

THERMOSTATED KAC MODELS

A Thesis
Presented to
The Academic Faculty

by

Ranjini Vaidyanathan

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Mathematics

Georgia Institute of Technology
December 2015

Copyright © 2015 by Ranjini Vaidyanathan

THERMOSTATED KAC MODELS

Approved by:

Dr. Federico Bonetto, Advisor
School of Mathematics
Georgia Institute of Technology

Prof. Rafael de la Llave
School of Mathematics
Georgia Institute of Technology

Prof. Evans Harrell
School of Mathematics
Georgia Institute of Technology

Prof. Michael Loss
School of Mathematics
Georgia Institute of Technology

Prof. Kurt Wiesenfeld
School of Physics
Georgia Institute of Technology

Date Approved: 30 October 2015

To my aunt and teacher,

Vasanthi Murali.

ACKNOWLEDGEMENTS

Foremost, I thank my advisor Dr. Federico Bonetto, and Prof. Michael Loss for their invaluable guidance and mentoring. I have truly enjoyed learning from, working with, and being inspired by them. I am especially thankful for the positive, encouraging environment they fostered in the group.

I thank my colleague and friend Hagop Tossounian for the stimulating mathematical discussions (and the witticisms during coffee breaks).

I thank Profs. Rafael de la Llave, Evans Harrell, and Kurt Wiesenfeld for being part of the thesis committee, and for their helpful comments.

I gratefully acknowledge the support from members of the School of Math throughout my studies at Georgia Tech. I am immensely thankful to my teachers, my family and my friends for their encouragement.

This dissertation work was partially supported by U.S. National Science Foundation grant DMS 1301555.

TABLE OF CONTENTS

DEDICATION	iii
ACKNOWLEDGEMENTS	iv
LIST OF FIGURES	vii
SUMMARY	viii
I INTRODUCTION	1
1.1 The Boltzmann Transport Equation and Propagation of Chaos . . .	3
1.2 Kac Model in Kinetic Theory	5
1.2.1 Approach to Equilibrium	10
1.3 Motivation	12
II THE INTERACTION OF THE KAC MODEL WITH A HEAT BATH	14
2.1 The Models	14
2.1.1 System Immersed in a Heat Bath	15
2.1.2 System in Partial Contact with a Heat Bath	16
2.2 List of Main Results	17
2.3 Approach to Equilibrium in L^2	24
2.3.1 Model I	25
2.3.2 Model II	34
2.4 Approach to Equilibrium in Entropy	40
2.4.1 Model I	40
2.4.2 Model II	51
2.5 Approach to Equilibrium in the Gabetta-Toscani-Wennberg Distance	59
2.6 Macroscopic Limit	62
2.7 Measures of Equilibration	65
III THE THERMOSTAT MODEL: TWO PERSPECTIVES	67
3.1 A Microscopic Realization	67

3.2 As a van Hove Limit	79
IV CONCLUSION AND OUTLOOK	82
APPENDIX A — SPECTRUM OF EVOLUTION OPERATOR FOR $N = 2, M = 1$	84
APPENDIX B — OPTIMIZING SEQUENCE FOR INEQUALITY (56)	86
REFERENCES	88

LIST OF FIGURES

1	van Hove Limit	80
---	--------------------------	----

SUMMARY

We consider a model of N particles interacting through a Kac-style collision process, with m particles among them interacting, in addition, with a thermostat. When $m = N$, we show exponential approach to the equilibrium canonical distribution in terms of the L^2 norm, in relative entropy, and in the Gabetta-Toscani-Wennberg (GTW) metric, at a rate independent of N . When $m < N$, the exponential rate of approach to equilibrium in L^2 is shown to behave as $\frac{m}{N}$ for N large, while the relative entropy and the GTW distance from equilibrium exhibit (at least) an “eventually exponential” decay, with a rate scaling as $\frac{m}{N^2}$ for large N . As an allied project, we obtain a rigorous microscopic description of the thermostat used, based on a model of a tagged particle colliding with an infinite gas in equilibrium at the thermostat temperature. These results are based on joint work with Federico Bonetto, Michael Loss and Hagop Tossounian.

CHAPTER I

INTRODUCTION

The primary theme of this dissertation is the question of approach to equilibrium in statistical mechanics. Statistical mechanics, which arose from kinetic theory, attempts to understand macroscopic properties of matter starting from the atomic hypothesis. Its aim is to unify the empirical laws of thermodynamics, hydrodynamics, etc. with the microscopic laws of physics obeyed by the constituents of matter.

A physical observable O (such as temperature) is associated with a function on the phase space $\Theta(\mathbf{x}, \mathbf{p})$ (correspondingly, kinetic energy) of the system, where $\mathbf{x}, \mathbf{p} \in \mathbb{R}^{3N}$ and N is the number of constituent particles. As the system evolves according to Newton's Laws (we only discuss the classical setting, where the energy scales are such that the laws of classical mechanics are a valid approximation), the function Θ is a highly fluctuating function of time. However, experimental measurements only sense average properties due to the presence of a large number of degrees of freedom. An appropriate limit is taken where $N \rightarrow \infty$, but physical properties like the density, kinetic energy per particle, etc. remain finite. This is called the thermodynamic limit.

The main postulates of statistical mechanics are that the state of a system can be fully described by a probability distribution on the phase space, and that in equilibrium (when macroscopic properties do not change), in the thermodynamic limit, the time average of an observable (the measured value) is equal to the phase space average of the corresponding function, with respect to an appropriately chosen probability distribution $f_{eq}(\mathbf{x}, \mathbf{p})$. That is,

$$\lim_{N \rightarrow \infty} \int d\mathbf{x} d\mathbf{p} \Theta(\mathbf{x}, \mathbf{p}) f_{eq}(\mathbf{x}, \mathbf{p}) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \Theta(\mathbf{x}(t), \mathbf{p}(t)) dt .$$

The above equation is called the ergodic hypothesis. Although proving that a

given mechanical system is ergodic is not easy to say the least, the setup for equilibrium statistical mechanics is well-defined. In particular, there is a prescription, formalized by Gibbs, for the choice of the probability distribution depending on the physical setting. Starting from this prescription, the laws of thermodynamics have been derived for many systems.

This is not the case when we are away from equilibrium, where the observables change on macroscopic time and space scales, and fluxes are involved. The general idea is that the probability distribution is now a function of time. There is not yet a general formalism that can be applied to all physical systems. In out-of-equilibrium statistical mechanics, we would like to answer on one hand, questions like how long a perturbed system takes to relax back to equilibrium. On the other hand, we would also want to investigate whether one can rigorously explain sustained non-equilibrium situations in which macroscopic measurements can be made, using ideas from the formalism of equilibrium statistical mechanics, like a local Gibbs distribution, and an appropriate limit of a large number of particles. In particular, we would like to understand what non-equilibrium steady state probability densities look like.

We mainly deal with the first question, but view our results as a step towards understanding the others. As is common in statistical mechanics, we use a stochastic system for our investigation instead of a deterministic one, for their better mixing properties (and hopefully similar statistical features); in particular, the Kac model of a spatially homogeneous dilute gas undergoing stochastic binary collisions, introduced by Mark Kac in 1956.

In this introductory chapter, we begin with a recap of the Boltzmann equation (a ground-breaking attempt to understand the notion of irreversibility starting from kinetic theory) and the Kac model in Sections 1.1 and 1.2. The motivation and scope of our work on the Kac model in a non-isolated setting is described in Section 1.3, where we also describe the outline of this thesis.

Apart from those explicitly cited in the text, the references consulted for this introductory chapter are [15, 4].

1.1 The Boltzmann Transport Equation and Propagation of Chaos

In the simplest possible case of a system of N identical, independent, non-interacting particles that move via Newton's Law, the evolution is essentially just streaming. If the initial state can be expressed by a product probability density $f(\mathbf{v}_1)\dots f(\mathbf{v}_N)$, then the product structure of the state is preserved by the evolution since the particles do not interact. The one-particle velocity distribution $f(\mathbf{v})$ is a function on \mathbb{R}^3 , and we consider the spatially homogeneous case so the probability distribution is independent of the spatial co-ordinate.

In 1872, Ludwig Boltzmann, in what he claimed was a theoretical justification of approach to equilibrium (The Second Law of Thermodynamics) from microscopic dynamics, published his (now) famous transport equation. He assumed that for a dilute gas in the macroscopic limit, particles behaved approximately independently and the state could be defined by a single-particle distribution. The molecules were modeled as hard spheres of radius R , moving according to classical mechanical laws and undergoing binary collisions among themselves. Their collisions with the boundary, and collisions among three or more particles were ignored.

He considered the limit of a large number of particles $N \rightarrow \infty$ in a finite volume V , but $R \rightarrow 0$ so that

- The gas remains dilute: the density $\rho \sim \frac{NR^3}{V}$ goes to 0.
- The collisions do not disappear in the limit: the number of collisions per unit time per particle, which is proportional to $\frac{N}{V}\bar{v}R^2$, where \bar{v} is the average speed, must remain finite and strictly positive in the limit.

$$R \rightarrow 0, N \rightarrow \infty, NR^3 \rightarrow 0, NR^2 \rightarrow \lambda, \tag{1}$$

where $\lambda > 0$ is finite.

Boltzmann made the assumption that under the above limit, called the Grad-Boltzmann limit, colliding particles always have uncorrelated pre-collision velocities. This allowed him to approximate the two-particle marginal $f_2(\mathbf{v}_1, \mathbf{v}_2)$ of the full N -particle velocity distribution $f(\mathbf{v}_1, \dots, \mathbf{v}_N)$ by a product of its single-particle marginals $f(\mathbf{v}_1)f(\mathbf{v}_2)$, and hence obtain an evolution equation for $f(\mathbf{v}_1)$ in terms of f alone. This assumption is called the *Stosszahlansatz*. Following is the spatially homogeneous Boltzmann equation for a gas of hard-spheres undergoing elastic collisions:

$$\frac{\partial f(\mathbf{v}, t)}{\partial t} = \int d\mathbf{n} \int d\mathbf{w} \{ f(\mathbf{w} + [\mathbf{n} \cdot (\mathbf{w} - \mathbf{v})]\mathbf{n}, t) f(\mathbf{w} - [\mathbf{n} \cdot (\mathbf{w} - \mathbf{v})]\mathbf{n}, t) - f(\mathbf{v}, t) f(\mathbf{w}, t) \} |(\mathbf{w} - \mathbf{v}) \cdot \mathbf{n}| , \quad (2)$$

where $\mathbf{n} \in S^2$, and $d\mathbf{n}$ is the corresponding surface area element. For this equation, Boltzmann introduced a Lyapunov functional, the entropy

$$H(t) := - \int f(\cdot, t) \log f(\cdot, t) , \quad (3)$$

and showed that it is non-decreasing in t , and that its maximum value is attained when f is taken to be the Maxwellian. Through this, Boltzmann showed approach to the equilibrium Maxwellian distribution $\left(\frac{\beta}{2\pi}\right)^{\frac{3}{2}} e^{-\frac{\beta|\mathbf{v}|^2}{2}}$ in the spatially homogeneous case.

The derivation of the equation presented was heuristic, and indeed, the Boltzmann equation took a long time to be accepted by the scientific community. Apart from doubts cast on validity of the assumptions made, the primary philosophical objections were that i) time-reversible microscopic dynamics could not lead to irreversible macroscopic behavior (Zermelo's paradox), and ii) the laws of mechanics mandate that the microscopic state always evolves to return infinitesimally close to the initial point, after a sufficiently long time interval (Poincaré recurrence).

The equation gained importance also because many physically interesting quantities can be expressed as (thermodynamic limits of) phase-space averages of functions

that depend only on a few co-ordinate/momenta variables, and hence the information contained in a small-particle marginal would suffice - this was the idea behind ‘Boltzmann statistical mechanics’. The Boltzmann equation is the cornerstone for macroscopic equations like the Euler and Navier-Stokes equations, and now has overwhelming empirical justification. The apparent paradoxes can be resolved: i) by interpreting the equation to describe close-to-equilibrium behavior (so the particles behave approximately independently and *Stosszahlansatz* would be a reasonable assumption), in an average sense, and in the presence of potential external noise and ii) by noting that the time scale of validity of the Boltzmann equation is much less than Poincaré recurrence times. Some of these counter-points were put forth by Boltzmann himself. In 1976, Oscar E. Lanford III [22, 23] gave a rigorous proof of eq. (2) for short times (of order less than the mean free flight time). In 1985, Illner and Pulvirenti [19] showed that the equation is valid for all times, but for the special case of a gas rapidly expanding into vacuum: one for which the *Stosszahlansatz* would hold true.

A different approach was taken by Mark Kac in 1956 to understand the essentials of Boltzmann’s heuristic derivation. He formulated a model that gained importance as a non-trivial system which, he showed, rigorously obeyed a Boltzmann Equation for all times - albeit one that stemmed from *stochastic* microscopic dynamics.

1.2 Kac Model in Kinetic Theory

Kac’s main motivation, as the title of his 1956 paper [21] says, was to understand kinetic theoretic foundations through a toy model for which Boltzmann’s equation is rigorously valid. He considered a system of N particles with 1-dimensional velocities (for simplicity) that interacted through stochastic binary collisions.

Let $\mathbf{v} = (v_1, \dots, v_N)$ represent the 1D velocities of the N particles. The Kac collision process can be described as follows: Pick a pair i, j uniformly among the particles

$1, \dots, N$. Pick an angle θ uniformly in $[0, 2\pi)$. Then, after a collision among particles i and j , the new state of the system becomes $\mathbf{v}_{i,j}(\theta) := (v_1, \dots, v_i^*, \dots, v_j^*, \dots, v_N)$, where

$$v_i^* = v_i \cos \theta + v_j \sin \theta$$

$$v_j^* = -v_i \sin \theta + v_j \cos \theta$$

The collision preserves the kinetic energy and hence $(v_1, \dots, v_N) \in S^{N-1}(\sqrt{2NE})$, the phase space, where E is the kinetic energy per particle. To ensure that the collision remains non-trivial, momentum conservation is not imposed.

Remark. Although the collision process is simple, it exhibits interesting behavior, and furthermore, paves the way to understanding its more physical momentum-preserving three-dimensional generalization - a Maxwellian gas [8].

Kac's idea was to replace the deterministic chaotic collisions, that were considered in the Boltzmann equation derivation, with random collisions that are Poisson distributed in time. The collisions occur at exponentially distributed time intervals with mean $\frac{1}{N\lambda}$, so that the collision rate is proportional to the number of particles N . The constant parameter $\lambda > 0$ is chosen through physical considerations, and is taken to be independent of time and of the state \mathbf{v} of the system. When the time for a collision event is reached, a choice i, j is made uniformly among the $\binom{N}{2}$ choices for pairs of particles, and for an angle $\theta \in [0, 2\pi)$ uniformly, and the velocities v_i, v_j are "rotated". Since we are interested in statistical properties, the primary object of interest is the probability distribution $f(\mathbf{v})$ on the phase space $S^{N-1}(\sqrt{2NE})$. We now proceed to derive an evolution equation for $f(\mathbf{v})$ based on the description of the above Markov jump process.

We first study the effect of the dynamics on a generic continuous function $\phi(\mathbf{v})$ on $S^{N-1}(\sqrt{2NE})$, and then translate this to its effect on the probability distribution $f(\mathbf{v})$. Initially, the expectation value of ϕ is

$$\int_{S^{N-1}} \phi(\mathbf{v}) f(\mathbf{v}, 0) .$$

The expectation value of $\phi(\mathbf{v})$ after a collision between particles i and j is

$$\int_{S^{N-1}} (R_{ij}\phi)(\mathbf{v}) f(\mathbf{v}, 0) d\mathbf{v} ,$$

where

$$R_{ij}\phi(\mathbf{v}) := \oint_0^{2\pi} \phi(\mathbf{v}_{i,j}(\theta)) d\theta$$

and $\oint_0^{2\pi}$ represents an averaging integral. Since the colliding particles are selected uniformly, the expectation after a generic collision event is

$$\int_{S^{N-1}} (Q\phi)(\mathbf{v}) f(\mathbf{v}, 0) ,$$

where $Q = \frac{1}{\binom{N}{2}} \sum_{i < j} R_{ij}$, the Markov transition operator, is called the Kac collision operator.

Due to the assumption of exponentially distributed wait times, we have that the number of collision events in a time interval $[0, t]$ is Poisson-distributed with intensity $N\lambda t$, and so the probability of k collision events occurring within this time is $e^{-N\lambda t} \frac{(N\lambda t)^k}{k!}$. Hence, the expectation of $\phi(\mathbf{v})$ at time t is:

$$\int_{S^{N-1}} e^{-N\lambda t} \sum_{k=0}^{\infty} \frac{N^k \lambda^k t^k}{k!} (Q^k \phi)(\mathbf{v}) f(\mathbf{v}, 0) .$$

Note that the term $Q^k \phi$ arises from the Markov property of the dynamics.

Finally, we use that ϕ is a generic function to shift the dynamics to the probability distribution (taking the adjoint), through

$$\int_{S^{N-1}} f(\mathbf{v}, t) \phi(\mathbf{v}) = \int_{S^{N-1}} \left\{ e^{-N\lambda t} \sum_{k=0}^{\infty} \frac{N^k \lambda^k t^k}{k!} (Q^k \phi)(\mathbf{v}) \right\} f(\mathbf{v}, 0) .$$

Then, the fact that $\int (R_{ij}\phi)(\mathbf{v}) f(\mathbf{v}) = \int \phi(\mathbf{v}) (R_{ij}f)(\mathbf{v})$, i.e., the rotation operator is self-adjoint on $L^2(S^{N-1})$, yields

$$f(\mathbf{v}, t) = e^{N\lambda(Q-I)t} f(\mathbf{v}, 0) .$$

It follows that the evolution of $f(\mathbf{v})$ is given by the master equation:

$$\frac{\partial f}{\partial t} = -\lambda N(I - Q)[f] . \tag{4}$$

Owing to indistinguishability among the particles, f is assumed to be symmetric in the v_i . This symmetry, and the normalization $\int f(\mathbf{v}) = 1$ is preserved by the evolution. The unique equilibrium is the uniform probability measure on $S^{N-1}(\sqrt{2NE})$, which follows from the fact that the Kac rotations R_{ij} generate any rotation on the sphere. As $N \rightarrow \infty$, number of particles explodes, and hence the total number of collisions explodes. However, the mean collision wait time for a specific particle is given by the mean wait time for *any* collision, $N\lambda$ times the probability that the specific particle is chosen, $\frac{2}{N}$. Analogous to Grad Boltzmann limit (1), the setup here is such that the mean collision wait time for a given particle (mean free time) remains finite and strictly positive independent of N .

Kac interpreted Boltzmann's molecular chaos hypothesis (*Stosszahlansatz*) in terms of finite-particle marginals of the phase space probability distribution. He introduced the notion of a chaotic state (Kac called it the 'Boltzmann property'), which is a weaker condition than that of a product probability distribution. The notion of a chaotic state formalizes the notion that particles are roughly independent in the thermodynamic limit.

Definition 1.2.1. [21] A sequence of probability densities $\{f^{(N)}(\mathbf{v})\}$ on $S^{N-1}(\sqrt{2NE})$ is said to be **chaotic** if $\forall k \geq 1$, and for arbitrary functions $\varphi_1, \dots, \varphi_k$ on \mathbb{R} ,

$$\lim_{N \rightarrow \infty} \int_{S^{N-1}(\sqrt{2NE})} f^{(N)}(\mathbf{v}) \varphi_1(v_1) \dots \varphi_k(v_k) = \prod_{j=1}^k \lim_{N \rightarrow \infty} \int_{S^{N-1}(\sqrt{2NE})} f^{(N)}(\mathbf{v}) \varphi_j(v_j) . \quad (5)$$

Kac showed in [21] that chaotic states are preserved by the evolution (4). This idea, now called "propagation of chaos", gives a rigorous probabilistic interpretation of *Stosszahlansatz*.

Theorem 1.2.2 (Propagation of Chaos). [21] *Let $f(\mathbf{v}, 0)$ be a chaotic state. Then $f(\mathbf{v}, t)$ as defined by eq. (4) is a chaotic state for every $t \geq 0$.*

Kac proved the above by expanding the exponential of the generator of the evolution (4), and showing that each term in the expansion on the left side of (5) converged

to the corresponding term on the right. He was able to interchange the summation and the limit as $N \rightarrow \infty$ for short times. Since chaotic states are preserved, he proved the result for all t by iterating the short-time result. In [26], McKean rewrote Kac's proof in a clear, algebraic manner.

This property of the Kac model was proved uniformly in time by Mischler, Mouhot [27] in 2013. Using Theorem 1.2.2, Kac showed that the single-particle density $\bar{f}(v)$ on \mathbb{R} , defined in a weak sense by

$$\int_{\mathbb{R}} \bar{f}(v_1) \varphi(v_1) = \lim_{N \rightarrow \infty} \int_{S^{N-1}(\sqrt{2NE})} f^{(N)}(\mathbf{v}) \varphi(v_1) , \quad (6)$$

satisfies the following Boltzmann equation in the thermodynamic limit.

$$\frac{\partial \bar{f}(v)}{\partial t} = 2\lambda \int d\theta \int dw [\bar{f}(v \cos \theta + w \sin \theta) \bar{f}(-v \sin \theta + w \cos \theta) - \bar{f}(v) \bar{f}(w)] . \quad (7)$$

An equilibrium for the above equation is one for which $f(v)f(w)$ is rotationally invariant as a function on \mathbb{R}^2 . This is reminiscent of Maxwell's elegant argument on how it is reasonable to impose that the velocity distribution $f(v_x, v_y, v_z)$ of an ideal gas in equilibrium i) does not depend on direction (rotationally invariant) and ii) has independence in the x, y, z co-ordinates of velocity (product), and that this implies that it is the Gaussian distribution. We now have that (7) has a family of equilibria $M_a := \frac{1}{\sqrt{2\pi a}} e^{-\frac{v^2}{2a}}$. However, the energy is fixed to be E by (6), and hence the equilibrium Maxwellian is M_E . Incidentally, it was Maxwell who first observed that the single-particle marginal of the uniform measure on the sphere in the $N \rightarrow \infty$ limit, is a Gaussian.

The Kac model was ideal for comparative studies between the single-particle Kac-Boltzmann equation and the N -particle, but *linear* master equation, as they were shown to be connected rigorously. In the thermodynamic limit, the evolution of a chaotic family of initial states $\{f^{(N)}(\mathbf{v})\}$ can be described equally accurately by the master equation as the non-linear Boltzmann equation.

1.2.1 Approach to Equilibrium

The Kac collision process is ergodic on the sphere, that is

$$\langle 2(I - Q)f, f \rangle = \frac{1}{\binom{N}{2}} \sum_{i < j} \int_{S^{N-1}} d\theta |f(\mathbf{v}_{i,j}(\theta)) - f(\mathbf{v})|^2 \geq 0 ,$$

where the inner product is in $L^2(S^{N-1})$, and $Qf = f \Leftrightarrow f = \frac{1}{|S^{N-1}(\sqrt{2NE})|}$. Kac conjectured [21] that any initial probability distribution $f(\mathbf{v}) \in L^2(S^{N-1})$ tends to the equilibrium uniform distribution exponentially at a rate strictly positive even as $N \rightarrow \infty$. This conjecture was proved by Janvresse [20], and shortly after, the spectral gap (the slowest exponential decay-rate in the L^2 norm) was found explicitly by Carlen, Carvalho, Loss [7] (see also [25]) using induction on the number of particles.

Theorem 1.2.3. [7] *The spectral gap of the Kac evolution operator,*

$$\inf\{\langle N(I - Q)f, f \rangle : f \in L^2(S^{N-1}), \|f\| = 1, \langle f, \mathbf{1} \rangle = 0\} = \frac{N+2}{2(N-1)} .$$

The gap eigenfunction is a fourth-degree spherical harmonic: $\sum_{j=1}^N v_j^4 - \frac{3}{N+2} (\sum_{j=1}^N v_j^2)^2$.

We thus have exponential approach to equilibrium in the L^2 metric:

$$\|f(\mathbf{v}, t) - \mathbf{1}\|_{L^2} \leq e^{-\frac{\lambda(N+2)}{2(N-1)}t} \|f(\mathbf{v}, 0) - \mathbf{1}\|_{L^2} ,$$

where $\mathbf{1}$ is the uniform probability measure on the sphere.

Remark. The appearance of the spherical harmonic as an eigenfunction is not surprising, since the Kac operator commutes with the Laplace-Beltrami operator on the S^{N-1} .

A more physical measure of equilibrium is the Gibbs entropy, the N -particle version of Boltzmann's entropy (3). Define

$$S(f) := \int_{S^{N-1}} f \log f .$$

We use the opposite sign, and hence approach to equilibrium in this case is indicated by the decrease in $S(f)$. It is easy to see, from convexity of $x \log x$, that $S(f) \geq 0$ and $S(f) = 0 \Leftrightarrow f = \mathbf{1}$. Furthermore, Kac showed in [21] that

$$\frac{dS}{dt} \leq 0$$

under the evolution (4).

A *quantitative* rate of decrease in entropy was found by Villani in 2003 [31]. He did so by showing that the entropy production $-\frac{dS}{dt}$ satisfies:

$$-\frac{dS(f(.,t))}{dt} \geq \frac{2\lambda}{N-1} S(f(.,t)) , \quad (8)$$

so that $S(f(.,t)) \rightarrow 0$ at an exponential rate inversely proportional to the number of particles. The result was obtained independently by Carlen, Lieb, Loss [5] using an induction argument.

If the bound is sharp, it would indicate that in the large-particle limit, a constant exponential decay-rate in entropy is not obtained. Indeed, for the Kac-Boltzmann equation (7), the entropy $S(f|M_E) = \int f \log \frac{f}{M_E}$ is measured relative to the equilibrium Gaussian density, M_E , and Cercignani's conjecture [10], which is

$$-\frac{dS(f(.,t)|M_E)}{dt} \geq k S(f(.,t)|M_E) \quad \text{for some } k > 0 ,$$

fails to hold. This was shown in [6], by explicitly constructing a sequence of states with finite, non-zero entropy but arbitrarily small entropy production. The sequence, similar to those used by [2], was a convex combination of Maxwellians, $(1-\delta)M_{\frac{E}{1-\delta}} + \delta M_{\frac{E}{\delta}}$. Each Maxwellian would be a solution to (7). The energy of the former tends to E while that of the latter goes to ∞ as $\delta \rightarrow 0$, but note that each portion contributes equally to the kinetic energy of the full distribution. Choosing δ arbitrarily close to 0, this becomes an optimizing sequence for the inequality $\frac{1}{S} \frac{dS}{dt} \leq 0$.

Thus, it was clear that entropic approach to equilibrium at an exponential rate uniform in N failed to persist in the thermodynamic limit, and the best result at

the level of the entropy production was the result $\frac{dS}{dt} \leq 0$ by Kac. Note that this doesn't preclude the possibility of a result involving higher order derivatives of the entropy, and one can speculate that the entropy decays exponentially after a slow, rate-determining initial decay period.

Returning to the many-particle master equation, Einav [13] showed that the entropy production bound (8) was essentially optimal by considering optimizing sequences similar to the above, but with δ depending on N . These were states in which a macroscopic fraction of the system's kinetic energy was contained in a vanishingly small fraction $N^{-\alpha}$ of the particles, for $\alpha > 0$ suitably chosen. One can wonder if restricting to states that we expect to see in nature would lead to better entropic approach to equilibrium (in particular, at an exponential rate uniform in N , so that it would persist in the thermodynamic limit). This leads us to the first problem considered in the thesis.

1.3 Motivation

One possibility to ensure proximity to “physical” states is to consider the Kac model in which each particle also interacts with a heat bath. Furthermore, it would be independently interesting to undertake a study of the Kac model in a non-isolated setting. This was the motivation for the first project [3], in collaboration with Federico Bonetto and Michael Loss. We study a system of N particles interacting via the Kac collision, and each of them also undergoing a Kac-type collision with a particle from a heat bath at inverse temperature β .

To remain close to the so-called physical states, is it necessary to apply the thermostat interaction to *all* particles? In the second project [30], which was work with Hagop Tossounian, we study equilibration in the Kac model with a proper subset of the particles connected to a heat bath at temperature $\frac{1}{\beta}$. It so turned out that in

the previous model, the equilibration persisted even in the absence of the Kac collision. Hence, this model helps us better understand the role of the Kac interaction in attaining equilibrium.

Another purpose was to introduce spatial inhomogeneity in the sense of identifying a subset of particles as situated “closer” to the heat bath. Hence, the thermostated particles would be the medium of transfer of energy and the role of the inter-particle interaction can be understood clearly. The rate-determining step is the transfer of energy between the thermostated and non-thermostated particles.

In both systems, we obtain results on the rate of approach to the equilibrium canonical distribution *quantitatively* using the L^2 distance, the relative entropy, and a metric related to the Fourier transform, and study its behavior in the thermodynamic limit. The results are detailed in Chapter 2.

In Chapter 3, we delve deeper into the thermostat model used, and show that it can be obtained as a large-size limit of non-ideal heat bath interactions. This work was in collaboration with Federico Bonetto, Michael Loss and Hagop Tossounian.

CHAPTER II

THE INTERACTION OF THE KAC MODEL WITH A HEAT BATH

In this chapter, we study a system of N particles interacting via the Kac collision, as before, but in a non-isolated setting where a portion of the system (m out of N particles) is allowed to interact with a heat bath. The kinetic energy of the system is not conserved.

The setting is non-trivial, and yet simple enough that a variety of physically interesting questions are mathematically tractable. In the following two sections, we explore in detail the two cases we consider: Model I where $m = N$, and Model II where $m < N$. They differ qualitatively from one another, as we shall see.

Most of the results described in this section have been published in [3, 30]. Sections 2.1 and 2.2 describe the models and list the main results. The remaining sections 2.3, 2.4, 2.5, 2.6 contain the proofs and discussion. We conclude in Section 2.7 by discussing the various measures of equilibration used in this Chapter.

2.1 *The Models*

As in the Kac model, spatial homogeneity is assumed, but the phase space changes to $\{\mathbf{v} \in \mathbb{R}^N\}$ as the heat bath dynamics will move the velocities out of the constant-energy sphere. The system is assumed to interact with an *ideal* heat bath whose particles remain in thermal equilibrium at temperature $\frac{1}{\beta}$; they are not affected by the interaction with the system. The heat bath interaction of, say Particle j , can be described through its action on a test function $\phi(v_1, \dots, v_N)$. Its transformation, given

this interaction, is

$$\int dw g(w) \oint d\theta \phi(\mathbf{v}_j(\theta, w)) =: W_j^* \phi ,$$

where $\mathbf{v}_j(\theta, w) = (v_1, \dots, v_j \cos \theta + w \sin \theta, \dots)$ and $g(w) = \sqrt{\frac{\beta}{2\pi}} e^{-\beta \frac{w^2}{2}}$ is the Maxwellian at the temperature of the heat bath. Hence, the interaction depicts a Kac style collision between the system particle j and a particle from an infinite gas in thermal equilibrium at temperature $\frac{1}{\beta}$. On a probability distribution $f(\mathbf{v})$ on the phase space \mathbb{R}^N , the action is given by the adjoint of the above operator, that is,

$$(W_j f)(\mathbf{v}) := \int dw \oint d\theta g(-v_j \sin \theta + w \cos \theta) f(\mathbf{v}_j(\theta, w)) . \quad (9)$$

We will refer to this operator as the **thermostat**.

One could consider a stronger operator than the above: one that ‘thermostats’ instantly. It forces Particle j to lose memory of its current distribution, and to acquire the Gaussian distribution at temperature $\frac{1}{\beta}$. On a test function $\phi(\mathbf{v})$, the result of Particle j ’s interaction is

$$\int dw_j g(w_j) \phi(\dots, w_j, \dots) =: S_j^* \phi ,$$

and on a probability density it becomes what we will call the **strong thermostat**:

$$S_j f(\mathbf{v}) = g(v_j) \int dw_j f(v_1, \dots, w_j, \dots, v_N) . \quad (10)$$

2.1.1 System Immersed in a Heat Bath

The system is modeled as a Markov jump process in which each of the N particles in the system interacts with the heat bath (as well as with each other). We will use the thermostat operators W_j from (9) to model the action of the heat bath. The collision events, i.e. the Kac and thermostat interaction, occur as a Poisson process with intensity $N(\lambda + \mu)$ with $\lambda, \mu > 0$ and proportional to the respective intensities of the Kac collision and the thermostat, and N is the number of particles. This scaling

is chosen, as in the Kac model, to ensure that the mean free time for a single particle is $O(1)$. The parameters λ and μ are chosen according to physical factors.

The wait times between “collision events” are distributed exponentially with mean $\frac{1}{N(\lambda+\mu)}$, and the transformation of a test function $\phi(\mathbf{v})$ after such an event is given by:

$$\left(\frac{\lambda}{\lambda + \mu} Q + \frac{\mu}{\lambda + \mu} \frac{1}{N} \sum_{i=1}^N W_j^* \right) \phi(\mathbf{v}) .$$

That is, once the time for an “event” is reached, the Kac collision is chosen with a probability $\frac{\lambda}{\lambda+\mu}$ and the thermostat interaction with probability $\frac{\mu}{\lambda+\mu}$, and within this, particles are chosen with uniform probability.

We now move to the probability distribution viewpoint and obtain the master equation for the evolution. The state of the system is described by the probability density $f(\mathbf{v}) \in L^1(\mathbb{R}^N)$. The reasoning is similar to the derivation of the Kac master equation (4), and f is seen to evolve through the equation:

$$\frac{\partial f}{\partial t} = -(\lambda \mathcal{G}_K + \mu \mathcal{G}_W) f , \quad (11)$$

where $\mathcal{G}_K := N(I - Q)$ is the Kac part, and $\mathcal{G}_W = \sum_{j=1}^N (I - W_j)$ corresponds to the thermostat that acts on each of the particles. Throughout, we assume symmetry among particles $1, \dots, N$, which is preserved by the evolution. The normalization $\int f(\mathbf{v}) = 1$ is also preserved. Henceforth, we will refer to this as **Model I**.

2.1.2 System in Partial Contact with a Heat Bath

Consider N particles labelled $1, \dots, N$ interacting through the Kac collisions, where $1, \dots, m$, with $m < N$, also interact with a heat bath at temperature $\frac{1}{\beta}$. To keep things simple, in this case we will use the strong thermostat operators S_j (see (10)) to model this interaction. The strong thermostat can be interpreted as an exchange of particles from the system with the heat bath. When particle j “hits” the heat bath, it is replaced by a particle from the heat bath that now assumes the label j . Since the heat bath particles have a Gaussian distribution, the effect is the eventual

equilibration of the whole system to the temperature $\frac{1}{\beta}$, where the Kac interaction plays a crucial role in the equilibration of the non-thermostated particles $(m+1, \dots, N)$.

We assume $m < N$, since the case $m = N$ has effectively been considered in the model described in the previous subsection (2.1.1). In this model, events occur at exponentially distributed times with mean $\frac{1}{N\lambda + m\mu}$, and such a collision event transforms a test function $\phi(\mathbf{v})$ as follows:

$$\left(\frac{N\lambda}{N\lambda + m\mu} Q + \frac{m\mu}{N\lambda + m\mu} \frac{1}{m} \sum_{j=1}^m S_j^* \right) \phi(\mathbf{v}) .$$

When the time for such an “event” is reached, the Kac collision or heat bath interaction is chosen, with probabilities proportional to $N\lambda$ and $m\mu$, respectively. The Markov transition operator is the adjoint of the above, and thus the master equation for the evolution of a probability density $f(\mathbf{v}) \in L^1(\mathbb{R}^N)$ becomes

$$\frac{\partial f}{\partial t} = -(\lambda \mathcal{G}_K + \mu \mathcal{G}_S) f , \quad (12)$$

where \mathcal{G}_K , the Kac part, acts among the N particles as before, while the thermostat operator $\mathcal{G}_S := \sum_{i=1}^m (I - S_i)$ acts only on Particles $1, \dots, m$. Here, we assume symmetry between $1, \dots, m$ and $m+1, \dots, N$, and this is preserved by the evolution. We will refer to this as **Model II**.

In the upcoming sections, we obtain quantitative results on the approach to equilibrium in the two systems described above, using various metrics to represent equilibration.

2.2 List of Main Results

In both the models considered, it is easy to see from eqs. (11) and (12) that the Gaussian

$$\gamma_\beta(\mathbf{v}) := \left(\sqrt{\frac{\beta}{2\pi}} \right)^N \prod_{i=1}^N e^{-\beta \frac{v_i^2}{2}}$$

is an equilibrium state. Moreover, it can be inferred from the results of section 2.3 that it is the unique equilibrium.

In contrast with the isolated Kac model (Section 1.2), the Kac operator \mathcal{G}_K now acts on functions on \mathbb{R}^N . Thus, when μ , the heat bath parameter is 0, both Models I and II have a degenerate steady state. Any function that is radial, i.e. depends only on $\|\mathbf{v}\|$, lies in the kernel of \mathcal{G}_K . The heat bath operators W and S , in Models I and II, are what determine *which* radial function the states evolve to: the above Gaussian.

The kinetic energy $K(t) := \frac{1}{2} \int_{\mathbb{R}^N} \left(\sum_{k=1}^N v_k^2 \right) f(\mathbf{v}, t) d\mathbf{v}$ of the equilibrium Gaussian state is $\frac{N}{2\beta}$, and we expect that both models have kinetic energies that approach this value. The Kac part \mathcal{G}_K preserves $\sum_k v_k^2$, and so below we study the effect of \mathcal{G}_W and \mathcal{G}_S .

From eq.(11), it is easy to obtain

$$\begin{aligned} \frac{dK}{dt} &= -\mu N K + \frac{\mu}{2} \sum_j \int dw d\mathbf{v} \oint d\theta \left(\sum_k v_k^2 \right) g(w_j^*(\theta)) f(\mathbf{v}_j(\theta, w)) \\ &= -\mu N K + \frac{\mu}{2} \sum_j \left(\sum_{k \neq j} \int d\mathbf{v} v_k^2 f(\mathbf{v}) + \int dw d\mathbf{v} g(w) f(\mathbf{v}) \oint d\theta (v_j \cos(\theta) + w \sin(\theta))^2 \right) \\ &= -\mu N K + \mu(N-1)K + \frac{\mu}{2} K + \frac{\mu}{2} \frac{N}{2\beta}. \end{aligned}$$

We thus get for Model I, that the kinetic energy approaches the equilibrium kinetic energy exponentially at a rate $\frac{\mu}{2}$.

$$\frac{dK}{dt} = -\frac{\mu}{2} \left(K - \frac{N}{2\beta} \right). \quad (13)$$

Setting

$$\frac{K(t)}{N} = \frac{1}{2} T(t), \quad (14)$$

the above equation reads as Newton's law of cooling when $T(t)$ is identified as the temperature at time t . However, this identification is not valid unless the probability density at time t is close to the Gaussian $\gamma_{\frac{1}{T(t)}}$. This is not true in general, except in the case of a quasi-static transformation, which can be achieved in Model I when $\mu \ll \lambda$, so that the Kac collisions are strong enough to effectively guide the evolution

through rotationally invariant states, whose finite-particle marginals are Gaussians in the thermodynamic limit.

Due to the asymmetry between particles $1, \dots, m$ and $m+1, \dots, N$ in Model II, we get a coupled system of differential equations for $K_m(t) := \frac{1}{2} \int_{\mathbb{R}^N} (\sum_{k=1}^m v_k^2) f(\mathbf{v}, t) d\mathbf{v}$ and $K_{N-m}(t) = \frac{1}{2} \int_{\mathbb{R}^N} (\sum_{k=m+1}^N v_k^2) f(\mathbf{v}, t) d\mathbf{v}$.

$$\begin{pmatrix} \dot{K}_{N-m} \\ \dot{K}_m \end{pmatrix} = \begin{pmatrix} -\frac{\lambda m}{N-1} & \frac{\lambda(N-m)}{N-1} \\ \frac{\lambda m}{N-1} & -\frac{\lambda(N-m)}{N-1} - \mu \end{pmatrix} \begin{pmatrix} K_{N-m} \\ K_m \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{m\mu}{2\beta} \end{pmatrix} \quad (15)$$

The equilibrium for the above is $\begin{pmatrix} \frac{N-m}{2\beta} \\ \frac{m}{2\beta} \end{pmatrix}$ and the smallest eigenvalue of the matrix (in magnitude) is

$$-\frac{1}{2} \left(\mu + \frac{N\lambda}{N-1} \right) \left(1 - \sqrt{1 - \frac{4m\lambda\mu}{N-1} \frac{1}{\left(\mu + \frac{N\lambda}{N-1} \right)^2}} \right).$$

Hence, the kinetic energy approaches the equilibrium value $\frac{N}{2\beta}$ at a rate $\sim \frac{m}{N}$ for N large.

Remark. We shall now set $\beta = 1$ without loss of generality.

Consider the transformation

$$f = \gamma(1 + h) \quad (16)$$

that defines h , the perturbation from the ground-state. Let h lie in the space $L^2(\mathbb{R}^N, \gamma(\mathbf{v}) d\mathbf{v})$ with inner product $\langle h_1, h_2 \rangle := \int h_1 h_2 \gamma d\mathbf{v}$. Under this transformation, the evolution operators in eqs. (11) and (12), transform into new operators that act on h , as $\mathcal{G}_K \longrightarrow \mathcal{L}_K$, $\mathcal{G}_W \longrightarrow \mathcal{L}_T$, and $\mathcal{G}_S \longrightarrow \mathcal{L}_P$. The operators are defined explicitly in Section 2.3. It then turns out that the operators corresponding to the time-evolution of h , $\mathcal{L} := \lambda \mathcal{L}_K + \mu \mathcal{L}_T$ for Model I, and $\mathcal{L}_{N,m} := \lambda \mathcal{L}_K + \mu \mathcal{L}_P$ for Model II, are both self-adjoint.

We have the following theorems on the lowest eigenvalues of the operator \mathcal{L} , and these indicate the rates of approach to equilibrium in Model I in terms of the $L^2(\mathbb{R}^N, \gamma)$ distance. The proofs are detailed in Section 2.3.

Theorem 2.2.1. *Define the spectral gap*

$$\Delta_N := \inf\{|\langle h, \mathcal{L}h \rangle| : \|h\| = 1, \langle h, 1 \rangle = 0\}.$$

We have

$$\Delta_N = \frac{\mu}{2}.$$

The corresponding eigenfunction is

$$h_{\Delta_N}(\mathbf{v}) := \sum_{i=1}^N (v_i^2 - 1).$$

Note that the parameter λ of the Kac operator does not appear in the gap, the slowest decay rate in the $L^2(\mathbb{R}^N, \gamma(\mathbf{v})d\mathbf{v})$ distance. To understand the role played by the Kac collision in the equilibration process, we define the “second” spectral gap

$$\Delta_N^{(2)} := \inf\{|\langle h, \mathcal{L}h \rangle| : \|h\| = 1, \langle h, 1 \rangle = 0, \langle h, h_{\Delta_N} \rangle = 0\},$$

which signifies the rate of decay when the initial distribution has no component in h_{Δ_N} . We have

Theorem 2.2.2. $\Delta_N^{(2)}$ *is given by the lower root of the quadratic equation*

$$x^2 - \left(\lambda\Lambda_N + \frac{13}{8}\mu\right)x + \mu\left(\lambda\Lambda_N + \frac{5}{8}\mu\right) - \frac{3}{8}\lambda\Lambda_N\mu\left(\frac{3}{N+2}\right) = 0, \quad (17)$$

where $\Lambda_N = \frac{1}{2}\frac{N+2}{N-1}$. *The corresponding eigenfunction is an even polynomial of degree 4 in all the v_i .*

As $N \rightarrow \infty$ one finds

$$\Delta_\infty^{(2)} = \min\left\{\frac{\lambda}{2} + \frac{5}{8}\mu, \mu\right\}. \quad (18)$$

Hence, in the large system limit, both the first and second spectral gaps are uniform in N . The proof of the above theorem uses the spectral gap of the Kac operator on the sphere (mentioned in Sec 1.2) to approximate the Kac operator \mathcal{L}_K with the orthogonal projection onto radial functions.

For Model II, the asymptotic behavior of the spectral gap

$$\delta_{N,m} := \inf\{|\langle h, \mathcal{L}_{N,m} h \rangle| : \|h\| = 1, \langle h, 1 \rangle = 0\} \quad (19)$$

for large N is given in the following theorem:

Theorem 2.2.3. *Assume $\lambda, \mu > 0$. Then for $1 \leq m \leq N - 1$,*

$$\frac{m}{N-1} \delta_{2,1} \leq \delta_{N,m} \leq \frac{m}{N-1} \frac{2\lambda\mu}{\mu + \lambda} . \quad (20)$$

where $\delta_{2,1} = \frac{(2\lambda+\mu)-\sqrt{4\lambda^2+\mu^2}}{2}$, which can be directly computed as it is an eigenvalue of a sum of projections (in the 2-particle case, the Kac operator $Q \equiv R_{12}$ is a projection). Thus, as we are close to equilibrium, $h \rightarrow 0$ in $L^2(\mathbb{R}^N, \gamma_\beta)$ at an exponential rate with a constant, which, when $m = \alpha N$, is uniform in the number of particles. In contrast with Model I, here the Kac collision plays a role in the equilibration of the slowest mode itself. This is expected, since $\lambda = 0$ means that the non-thermostated particles $m + 1, \dots, N$ never reach the temperature of the heat bath.

In Section 2.4, we take up a study of the relative entropy of a density f with respect to the equilibrium Gaussian $\gamma(\mathbf{v})$ defined as

$$S(f|\gamma) = \int f \log \frac{f}{\gamma} d\mathbf{v} .$$

This functional is strictly positive except at equilibrium, i.e $S(\gamma|\gamma) = 0$. One could try to prove Cercignani's conjecture [10] for the Models by finding a bound for the entropy production $-\frac{dS}{dt}$, of the form

$$-\frac{dS(f_t|\gamma)}{dt} \geq kS(f_t|\gamma) \quad (21)$$

for some $k > 0$, which implies the exponential bound $S(f(.,t)|\gamma) \leq e^{-kt}S(f(.,0)|\gamma)$ for the entropy. The left-hand side of (21) can be written in terms of the generator of the evolution and its properties could be used to find a bound. In both of our systems, however, the better approach seems to be to employ convexity of the entropy to expand the term $S(e^{-(\lambda\mathcal{G}_K+\mu\mathcal{G}_W)t}f|\gamma)$, and $S(e^{-(\lambda\mathcal{G}_K+\mu\mathcal{G}_S)t}f|\gamma)$ for Models I and II, respectively, and find an upper bound for the rate of decay.

Remark. This approach of using the entropy directly was used in [5] to obtain the entropy decay bound (8) for the isolated Kac model.

We have the following Theorem for Model I:

Theorem 2.2.4. *Let $f(\mathbf{v},t)$ be the solution of the master equation (11) with initial condition $f(\mathbf{v},0)$. Then*

$$S(f(.,t)|\gamma) \leq e^{-\rho t}S(f(.,0)|\gamma) ,$$

where

$$\rho = \frac{\mu}{2} .$$

We prove this by first considering a single-particle system ($N = 1$), where we relate the thermostat W_1 on one particle to the Ornstein-Uhlenbeck semigroup, and use the entropy decay properties of the latter [28, 16, 1, 17, 29]. To generalize to higher values of N , we invoke Han's inequality [18] in information theory (this can also be inferred from the Loomis-Whitney inequality [24]), which bounds the sum of entropies of marginal distributions in terms of the total entropy.

Theorem 2.2.4 implies an entropy production bound like (21): $\frac{dS}{dt} \leq -\frac{\mu}{2}S$. The strength of the Kac term λ does not play a role in this bound since the entropy production due to the Kac term is zero on any radial function. One may wonder if we can obtain a better entropy production bound by taking into account the coupling between the Kac and thermostat terms, but it turns out that $k = \frac{\mu}{2}$ is optimal, and

will be shown by explicitly constructing an optimizing sequence along the lines of [2, 6, 13].

For Model II, we have the following entropy decay property:

Theorem 2.2.5 ([30]). *Assume $1 \leq m < N$ and let $f(\mathbf{v}, t)$ be the solution of the master equation (12) with initial condition $f(\mathbf{v}, 0)$. Then*

$$S(f(., t)|\gamma) \leq \left(-\frac{\xi_-}{\xi_+ - \xi_-} e^{-\xi_+ t} + \frac{\xi_+}{\xi_+ - \xi_-} e^{-\xi_- t} \right) S(f(., 0)|\gamma), \quad (22)$$

where $\xi_{\pm} \equiv \xi_{\pm}(m, N) = \left(\frac{N\lambda + \mu}{2} \pm \frac{1}{2} \sqrt{(N\lambda + \mu)^2 - 4m\lambda\mu/(N-1)} \right)$.

Define

$$Z(t) := -\frac{\xi_-}{\xi_+ - \xi_-} e^{-\xi_+ t} + \frac{\xi_+}{\xi_+ - \xi_-} e^{-\xi_- t}.$$

For $\lambda, \mu > 0$, $\lim_{t \rightarrow \infty} Z(t) = 0$ and $Z(0) = 1$. In fact, $Z'(t) \leq 0$ since $\xi_- < \xi_+$. The slowest decay mode (the dominant rate as $t \rightarrow \infty$) is the one corresponding to $e^{-\xi_- t}$, and for large N , $\xi_- \sim \frac{m\lambda\mu}{(N-1)(N\lambda+\mu)}$. Hence, we obtain an eventually exponential decay of relative entropy through this bound, albeit with decay constant $\sim \frac{m}{N^2}$. Unfortunately, since $Z'(0) = 0$, the Theorem does not give us a bound of type (21) on the entropy production.

Next, in Section 2.5, we consider a metric introduced by Gabetta, Toscani, Wennberg [14] that is related to the Fourier transform of probability distributions. We show in Theorem 2.5.1 that this metric, when used to measure the distance from equilibrium, behaves exactly like the relative entropy. The results in this section were obtained later, and hence do not appear in [3, 30].

In Section 2.6, we examine the notion of propagation of chaos applied to Models I and II. The main result is Theorem 2.6.2, which shows propagation of chaos for Model I, which enables one to connect it rigorously with a Boltzmann-type equation. If we define

$$\bar{f}_t(v_1) = \lim_{N \rightarrow \infty} \int f_t^{(N)}(\mathbf{v}) dv_2 \cdots dv_N,$$

then eq. (11), through Theorem 2.6.2, gives rise to an effective evolution for \bar{f}_t , that is

Theorem 2.2.6. $\bar{f}_t(v)$ is the solution of the following ‘Boltzmann Equation’:

$$\begin{aligned} \frac{\partial \bar{f}_t(v)}{\partial t} = & 2\lambda \oint d\theta \int dw [\bar{f}_t(v \cos \theta + w \sin \theta) \bar{f}_t(-v \sin \theta + w \cos \theta) - \bar{f}_t(v) \bar{f}_t(w)] \\ & + \mu \left[\int dw \oint d\theta g(-v \sin \theta + w \cos \theta) \bar{f}_t(v \cos \theta + w \sin \theta) - \bar{f}_t(v) \right] \end{aligned}$$

with $\bar{f}_0(v)$ as initial condition.

Our proof follows the one for the isolated Kac model in [21, 26], and hence it does not establish the validity of the above Boltzmann equation *uniformly* in time. We believe it should be possible to prove such a result by adapting the proof in [27] where propagation of chaos for the isolated Kac model is shown uniformly in time. The additional thermostat term in our case preserves product states, and hence should not cause a hindrance in adapting the proof.

2.3 Approach to Equilibrium in L^2

As in the case of the isolated Kac model, the question of equilibration in Models I and II can be translated to one in spectral theory, through the ground-state transformation (16) that leads to self-adjoint evolution operators. We can then focus on the spectral properties of these operators - in particular the smallest eigenvalues (in magnitude) - and infer from this the equilibration rate of the system. Note however, that in doing this, we are restricting to distributions f such that $h = \frac{f}{\gamma} - 1$ is in $L^2(\mathbb{R}^N, \gamma)$. For instance, Gaussian states with temperature greater than twice that of the heat bath do not satisfy this constraint.

In concurrence with the ground-state transformation (16), let $\mathcal{X}_N := \{h \in L^2(\mathbb{R}^N, \gamma) : \langle h, 1 \rangle = 0\}$, where $\langle \cdot, \cdot \rangle$ denotes the inner product in the L^2 space with weight γ . The condition $\langle h, 1 \rangle = \int h(\mathbf{v}) \gamma(\mathbf{v}) d\mathbf{v} = 0$ corresponds to the normalization of the probability density f . In the following two subsections, we address the rate at which $h \rightarrow 0$,

the equilibrium, in the Hilbert space \mathcal{X}_N for each of the models above. In particular, we prove Theorems 2.2.1, 2.2.2 and 2.2.3.

2.3.1 Model I

The master equation (11) written in terms of h is

$$\frac{\partial h}{\partial t} = -(\lambda \mathcal{L}_K + \mu \mathcal{L}_T)h, \quad (23)$$

where $\mathcal{L}_K := N(I - Q)$ corresponds to the Kac collisions and $\mathcal{L}_T := \sum_{j=1}^N (I - T_j)$ to the thermostat, and

$$T_j h := \oint d\theta \int dw g(w) h(\mathbf{v}_j(\theta, w)). \quad (24)$$

The Kac operator does not change under the ground-state transformation, i.e., $\mathcal{L}_K = \mathcal{G}_K$, whereas the thermostat does. Recall that the vector $\mathbf{v}_j(\theta, w) = (v_1, \dots, v_j \cos \theta + w \sin \theta, \dots)$ is the velocity vector obtained following the Kac collision of particle j with a thermostat particle with velocity w .

It is easy to see that the operator $\mathcal{L} := \lambda \mathcal{L}_K + \mu \mathcal{L}_T$ for the evolution of h is self-adjoint on \mathcal{X}_N . Moreover \mathcal{L} preserves the subspace of \mathcal{X}_N formed by the functions symmetric under permutation of the variables.

To begin, we report some known or simple results on the spectra of \mathcal{L}_K and \mathcal{L}_T . We say that a function $h(\mathbf{v})$ is *radial* if it depends only on $r^2 = \sum_i v_i^2$. We call \mathcal{X}_r the subspace of \mathcal{X}_N that consists of radial functions. We have

Lemma 2.3.1.

- $\mathcal{L}_K \geq 0, \mathcal{L}_T \geq 0$.
- $\mathcal{L}_K[h] = 0 \Leftrightarrow h \in \mathcal{X}_r$, and $\mathcal{L}_T[h] = 0 \Leftrightarrow h = \text{constant}$.

Proof. All claims follow from the following observations:

$$2\langle (I - Q)h, h \rangle = \frac{1}{\binom{N}{2}} \sum_{i < j} \int d\theta \int_{\mathbb{R}^N} |h(\mathbf{v}_{i,j}(\theta)) - h(\mathbf{v})|^2 \gamma d\mathbf{v} \geq 0$$

$$2\langle \sum_j (I - T_j)h, h \rangle = \sum_j \left(\int d\theta \int d\mathbf{v} dw g(w) \gamma(\mathbf{v}) |h(\mathbf{v}_j(\theta, w)) - h(\mathbf{v})|^2 \right) \geq 0 ,$$

the first of which is an identity due to Kac [21]. \square

Notice that the Kac operator alone acting on \mathbb{R}^N has a degenerate ground state. From the above Lemma, we see that the unique equilibrium state in \mathcal{X}_N corresponding to eq. (23) is $h(\mathbf{v}) = 0$.

The following Theorem is a direct consequence of the results in [7], and is equivalent to Theorem 1.2.3.

Theorem 2.3.2 ([7]). *We have that*

$$\Lambda_N := \inf\{|\langle h, \mathcal{L}_K h \rangle| : \|h\| = 1, h \perp \mathcal{X}_r\} = \frac{1}{2} \frac{N+2}{N-1}$$

and the corresponding eigenfunction is $\sum_{j=1}^N v_j^4 - \frac{3}{N+2} \left(\sum_{j=1}^N v_j^2 \right)^2$.

To study the spectrum of \mathcal{L}_T we use the Hermite polynomials $H_\alpha(v)$ with weight $g(v)$. More precisely, for α integer, we set

$$H_\alpha(v) = (-1)^\alpha e^{\frac{v^2}{2}} \frac{d^\alpha}{dv^\alpha} e^{-\frac{v^2}{2}}$$

so that

1. $H_\alpha(v)$ is a polynomial of degree α . Moreover $H_\alpha(-v) = (-1)^\alpha H_\alpha(v)$.
2. The coefficient of v^α in H_α is 1.
3. The H_α are orthogonal in $L^2(\mathbb{R}, g dv)$. More precisely

$$\int H_{\alpha_1}(v) g(v) H_{\alpha_2}(v) dv = \sqrt{2\pi} \alpha_1! \delta_{\alpha_1, \alpha_2}.$$

Lemma 2.3.3. $H_\alpha(v_j)$ form an orthogonal basis of eigenfunctions for the operator T_j and $T_j H_\alpha = s_\alpha H_\alpha$ with $s_\alpha = 0$ if α is odd while

$$s_{2\alpha} = \oint_0^{2\pi} d\theta \cos^{2\alpha} \theta = \frac{(2\alpha)!}{2^{2\alpha} \alpha!^2} .$$

Proof. We drop the subscript j here for ease of notation. First, we observe that

$$\begin{aligned} \int T[H_\alpha(v)] H_n(v) g(v) dv &= \int dw dv g(v) g(w) H_n(v) \oint d\theta H_\alpha(v \cos \theta + w \sin \theta) \\ &= \int dw dv g(v) g(w) H_\alpha(v) \oint d\theta H_n(v \cos \theta + w \sin \theta). \end{aligned}$$

(by self-adjointness)

As T preserves polynomial degrees, $T[H_\alpha(v)]$ is a polynomial in v of degree α , and so the first line implies that $\int T[H_\alpha(v)] H_n(v) g(v) dv = 0$ if $n > \alpha$. Likewise, the second line implies that $\int T[H_\alpha(v)] H_n(v) g(v) dv = 0$ if $\alpha > n$. Thus,

$$T[H_\alpha(v)] = c_\alpha H_\alpha(v) .$$

By equating the coefficients of v^α in the above, we get that $c_\alpha = \oint \cos^\alpha \theta = s_\alpha$. \square

Note that $s_{2(\alpha+1)} < s_{2\alpha}$ and $s_{2\alpha} \rightarrow 0$ as $\alpha \rightarrow \infty$. Since \mathcal{L}_T is just the direct sum of $(I - T_j)$ we get the following characterization of the spectrum of \mathcal{L}_T .

Corollary 2.3.4. *The functions*

$$H_{\underline{\alpha}}(\mathbf{v}) := \prod_{i=1}^N H_{\alpha_i}(v_i) ,$$

where $\underline{\alpha} = (\alpha_1, \dots, \alpha_N)$, is an eigenfunction of \mathcal{L}_T with eigenvalue

$$\sigma_{\underline{\alpha}} := \sum_i (1 - s_{\alpha_i}).$$

The set $\{H_{\underline{\alpha}}\}_{\underline{\alpha} \geq 0}$ form an orthogonal basis of eigenfunctions for \mathcal{L}_T in \mathcal{X} . In particular $\mathcal{L}_T > 0$ on \mathcal{X}^\perp .

To study the spectral gap, we need to understand the action of \mathcal{L}_K on products of Hermite polynomials $H_{\underline{\alpha}}$, the eigenfunctions of \mathcal{L}_T . We first state and prove the following lemma which helps us restrict our investigation to even polynomials.

Lemma 2.3.5.

- Any eigenfunction of $\mu\mathcal{L}_T + \lambda\mathcal{L}_K$ is either even or odd in each variable v_i .
- If E is an eigenvalue of $\mu\mathcal{L}_T + \lambda\mathcal{L}_K$, with an eigenfunction that is odd in some v_i , we have that $E \geq 2\lambda + \mu$.

Proof. The first part can be seen by noting that the operator $\mu\mathcal{L}_T + \lambda\mathcal{L}_K$ commutes with the reflection operator $S_j[h](\mathbf{v}) := h(\dots, -v_j, \dots)$. For the second part, say $(\mu\mathcal{L}_T + \lambda\mathcal{L}_K)h = Eh$, with $S_1[h] = -h$. Then $T_1[h] = 0$. In addition, for any $i \neq 1$,

$$\begin{aligned}
\oint d\theta h(\mathbf{v}_{i,1}(\theta)) &= \oint d\theta h(v_i \cos \theta + v_1 \sin \theta, \dots, -v_i \sin \theta + v_1 \cos \theta, \dots) \\
&= \oint d\theta h(\sqrt{v_i^2 + v_1^2} \cos(\varphi - \theta), \dots, \sqrt{v_i^2 + v_1^2} \sin(\varphi - \theta), \dots) \\
&= \oint d\theta h(\sqrt{v_i^2 + v_1^2} \cos \theta, \dots, \sqrt{v_i^2 + v_1^2} \sin \theta, \dots) \\
&= \oint d\theta h(-\sqrt{v_i^2 + v_1^2} \cos \theta, \dots, \sqrt{v_i^2 + v_1^2} \sin \theta, \dots) \quad (\text{taking } \theta \rightarrow \pi - \theta) \\
&= 0 \quad (\text{using that } S_1[h] = -h).
\end{aligned}$$

Thus,

$$\lambda N h - \lambda \frac{N}{\binom{N}{2}} \sum_{i < j, i, j \neq 1} \oint d\theta h(\mathbf{v}_{i,j}(\theta)) + N\mu h - \mu \sum_{i \neq 1} T_i[h] = Eh$$

or

$$(\lambda N + \mu N - E) \leq \lambda \frac{N}{\binom{N}{2}} \binom{N-1}{2} + \mu(N-1),$$

which proves the claim. \square

We will show that the eigenfunctions of interest have corresponding eigenvalues that are smaller than $2\lambda + \mu$. We can thus restrict our attention to the space of functions that are even in all variables. To this end we define

$$L_{2l} = \text{span} \left\{ H_{2\alpha} \mid \sum_{i=1}^N 2\alpha_i = 2l \right\}.$$

Moreover we set

$$|\underline{\alpha}| = \sum_{i=1}^N \alpha_i$$

and

$\Xi := \{\underline{\alpha} : \sum_{i < j} \alpha_i \alpha_j \neq 0\}$, that is the set of $\underline{\alpha}$ in which at least two entries are non-zero.

Lemma 2.3.6. *In each L_{2l} the eigenvalues of \mathcal{L}_T are given by $\sigma_{2\underline{\alpha}} = \sum_j (1 - s_{2\underline{\alpha}_j})$, where $|\underline{\alpha}| = l$. It follows that*

- *The smallest eigenvalue in each L_{2l} is $1 - s_{2l}$ and the corresponding eigenfunctions are precisely linear combinations of $H_{2\underline{\alpha}}(\mathbf{v})$ with $\underline{\alpha} = (0, \dots, l, \dots, 0)$.*
- $\min_{\underline{\alpha} \in \Xi} \sigma_{2\underline{\alpha}} = 1$. *Moreover, the minimum is reached when two of the α_i 's are 1 and the rest are 0.*

Proof. To prove the first statement, we start by observing that the function $J(x) := \int_0^{2\pi} \cos^{2x} \theta d\theta$ is strictly convex in x . Consider $\underline{\alpha}$ such that $|\underline{\alpha}| = l$. We need to show that

$$\sum J(\alpha_i) \leq J(l) + (N - 1)J(0)$$

and that equality is attained if and only if $\underline{\alpha} = (0, \dots, l, \dots, 0)$. By convexity, we have that

$$\begin{aligned} J(\alpha_i) &= J\left(\frac{\alpha_i}{l}l + \sum_{j \neq i} \frac{\alpha_j}{l}0\right) \\ &\leq \frac{\alpha_i}{l}J(l) + \sum_{j \neq i} \frac{\alpha_j}{l}J(0) . \end{aligned}$$

Summing the above over i , we get the result.

The second claim follows from the monotonicity of the $s_{2\underline{\alpha}}$ and the fact that $s_2 = \frac{1}{2}$. □

We now have all the ingredients to find the spectral gap of \mathcal{L} .

Proof of Theorem 2.2.1. By Corollary 2.3.4 and Lemma 2.3.6, we have that $\mathcal{L}_T \geq 1/2$ and thus $\mathcal{L} \geq 1/2$ on \mathcal{X}_N . On the other hand, $\mathcal{L}[\sum H_2(v_i)] = \mu \mathcal{L}_T[\sum H_2(v_i)] =$

$\frac{\mu}{2}(\sum H_2(v_i))$ since $\sum H_2(v_i)$, being a radial function, is annihilated by the Kac part. Thus, $\Delta_N = \mu/2$ and $h_{\Delta_N} = \sum_{i=1}^N H_2(v_i) \in L_2$. \square

To compute $\Delta_N^{(2)}$ we need to better understand the action of \mathcal{L}_K on the L_{2l} . This is done in the following Lemma, which is actually a generalization of Lemma 2.3.3.

Lemma 2.3.7. *Let A be a self-adjoint operator on $L^2(\mathbb{R}^N, \gamma(\mathbf{v})d\mathbf{v})$ that preserves the space P_{2l} , of homogeneous even polynomials in v_1, \dots, v_N of degree $2l$. If*

$$A(v_1^{2\alpha_1} \dots v_N^{2\alpha_N}) = \sum_{|\underline{\beta}|=|\underline{\alpha}|} c_{\underline{\beta}} v_1^{2\beta_1} \dots v_N^{2\beta_N} ,$$

we get

$$A(H_{2\alpha_1}(v_1) \dots H_{2\alpha_N}(v_N)) = \sum_{|\underline{\beta}|=|\underline{\alpha}|} c_{\underline{\beta}} H_{2\beta_1}(v_1) \dots H_{2\beta_N}(v_N) .$$

Proof. First, we observe that $A(L_{2l}) \subset L_{2l}$. Indeed, if $f \in L_{2m}$ and $g \in L_{2l}$ with $m < l$, we have $\langle Ag, f \rangle = \langle g, Af \rangle = 0$ because Af contains only monomials of degree at most $2m$. This means that

$$A(H_{2\alpha_1}(v_1) \dots H_{2\alpha_N}(v_N)) = \sum_{|\underline{\beta}|=|\underline{\alpha}|} k_{\underline{\beta}} H_{2\beta_1}(v_1) \dots H_{2\beta_N}(v_N)$$

and because

$$A(v_1^{2\alpha_1} \dots v_N^{2\alpha_N}) = \sum_{|\underline{\beta}|=|\underline{\alpha}|} c_{\underline{\beta}} v_1^{2\beta_1} \dots v_N^{2\beta_N} ,$$

we get that $c_{\underline{\beta}} = k_{\underline{\beta}}$ for any $\underline{\beta}$ by equating the coefficients of the term of maximal degree $v_1^{2\beta_1} \dots v_N^{2\beta_N}$. \square

Remarks.

- Since \mathcal{L}_K preserves the spaces P_{2l} , the above Lemma applies to it. Thus, the action of \mathcal{L}_K on products of Hermite polynomials $H_{2n}(v_i)$ can be deduced from its action on products of monomials v_i^{2n} , and the latter turns out to be simpler.
- Note that L_{2l} is invariant under \mathcal{L}_K and thus is invariant under \mathcal{L} .

In preparation for computing the “second spectral gap” $\Delta_N^{(2)}$ we note that Theorem 2.3.2 implies that

$$\langle h, \mathcal{L}_K h \rangle \geq \langle h, \Lambda_N(I - B)h \rangle ,$$

where B is the orthogonal projection on radial functions, that is

$$B[h](\mathbf{v}) = \int_{S^{N-1}(|\mathbf{v}|)} h(\mathbf{w}) d\sigma(\mathbf{w}) . \quad (25)$$

where $S^{N-1}(r)$ is the sphere of radius r in \mathbb{R}^N with normalized surface measure $d\sigma(\mathbf{v})$.

Setting $\mathcal{L}_R := \Lambda_N(I - B)$ we have

$$\langle h, \mathcal{L}h \rangle \geq \langle h, (\mu\mathcal{L}_T + \lambda\mathcal{L}_R)h \rangle$$

so that

$$\Delta_N^{(2)} \geq \inf\{\langle h, (\mu\mathcal{L}_T + \lambda\mathcal{L}_R)h \rangle : \|h\| = 1, h \perp L_0, L_2\} , \quad (26)$$

where we have replaced the operator \mathcal{L}_K with the much simpler projection \mathcal{L}_R . Note, the same reasoning as before shows that the space L_{2l} is invariant under \mathcal{L}_R . For later use we define

$$\Gamma(\underline{\alpha}) = \int_{S^{N-1}(1)} v_1^{2\alpha_1} \dots v_N^{2\alpha_N} d\sigma_1(\mathbf{v}).$$

Theorem 2.3.8. *The smallest eigenvalue a_l of the operator*

$$\mu\mathcal{L}_T + \lambda\mathcal{L}_R$$

restricted to the space L_{2l} satisfies the estimates

$$a_l \geq x_l ,$$

where x_l is the smaller of the two solutions of the equation

$$x^2 - (\lambda\Lambda_N + (2 - s_{2l})\mu)x + (1 - s_{2l})\mu^2 + \lambda\Lambda_N\mu = \lambda\Lambda_N\mu s_{2l}N\Gamma(l, 0, \dots, 0) . \quad (27)$$

Proof. The equation for the eigenvalue x of $\mu\mathcal{L}_T + \lambda\mathcal{L}_R$ gives

$$\mu \sum T_j h + \lambda\Lambda_N B h = (N\mu + \lambda\Lambda_N - x)h .$$

Observe that if $|\underline{\alpha}| = l$, $B[v_1^{2\alpha_1} \dots v_N^{2\alpha_N}]$ is an homogeneous radial polynomial of degree $2l$ so that we have

$$B[v_1^{2\alpha_1} \dots v_N^{2\alpha_N}](r) = \Gamma(\underline{\alpha}) r^{2l} = \Gamma(\underline{\alpha}) \sum_{|\underline{\beta}|=l} \frac{l!}{\beta_1! \dots \beta_N!} v_1^{2\beta_1} \dots v_N^{2\beta_N}, \quad (28)$$

in particular

$$\sum_{|\underline{\alpha}|=l} \frac{l!}{\alpha_1! \dots \alpha_N!} \Gamma(\underline{\alpha}) = 1. \quad (29)$$

Writing a generic function f in L_{2l} as

$$f = \sum_{|\underline{\alpha}|=l} c_{\underline{\alpha}} H_{2\underline{\alpha}}$$

the eigenvalue equation becomes:

$$\mu \sum_{|\underline{\alpha}|=l} \sum_j s_{2\alpha_j} c_{\underline{\alpha}} H_{2\underline{\alpha}} + \lambda \Lambda_N \left[\sum_{|\underline{\alpha}|=l} c_{\underline{\alpha}} \Gamma(\underline{\alpha}) \right] \sum_{|\underline{\alpha}|=l} \frac{l!}{\alpha_1! \dots \alpha_N!} H_{2\underline{\alpha}} = (N\mu + \lambda \Lambda_N - x) \sum_{|\underline{\alpha}|=l} c_{\underline{\alpha}} H_{2\underline{\alpha}}, \quad (30)$$

where we have used that the projection B satisfies the hypothesis of Lemma 2.3.7.

Thus for every $\underline{\alpha}$

$$(\mu \sigma_{2\underline{\alpha}} + \lambda \Lambda_N - x) c_{\underline{\alpha}} = K \lambda \Lambda_N \frac{l!}{\alpha_1! \dots \alpha_N!}, \quad (31)$$

where we set $\sum_{|\underline{\alpha}|=l} c_{\underline{\alpha}} \Gamma(\underline{\alpha}) = K$. Consider first the case $K \neq 0$, that is $(x - \lambda \Lambda_N - \mu \sigma_{2\underline{\alpha}}) \neq 0$ for every $\underline{\alpha}$. Rearranging, multiplying both sides by $\Gamma(\underline{\alpha})$, and adding we get

$$\frac{1}{\lambda \Lambda_N} = \sum_{|\underline{\alpha}|=l} \frac{1}{\lambda \Lambda_N + \mu \sigma_{2\underline{\alpha}} - x} \Gamma(\underline{\alpha}) \frac{l!}{\alpha_1! \dots \alpha_N!}. \quad (32)$$

With x moving in from $-\infty$, the first singularity of the right side of eq. (32) occurs when

$$x = \min_{|\underline{\alpha}|=l} (\lambda \Lambda_N + \mu \sigma_{2\underline{\alpha}}) = \lambda \Lambda_N + \mu(1 - s_{2l}),$$

where the last equality follows from Lemma 2.3.6. The right side of eq. (32) is a positive increasing function of x until the first singularity. Thus, the smallest

eigenvalue is less than $\lambda\Lambda_N + \mu(1 - s_{2l})$. For $0 < x < \lambda\Lambda_N + \mu(1 - s_{2l})$ we get

$$\begin{aligned}
\frac{1}{\lambda\Lambda_N} &= \frac{1}{\lambda\Lambda_N + (1 - s_{2l})\mu - x} N\Gamma(l, 0, \dots, 0) + \sum_{\substack{|\underline{\alpha}|=l \\ \underline{\alpha} \in \Xi}} \frac{1}{(\lambda\Lambda_N + \mu\sigma_{2\underline{\alpha}} - x)} \Gamma(\underline{\alpha}) \frac{l!}{\alpha_1! \dots \alpha_N!} \\
&\leq \frac{1}{\lambda\Lambda_N + (1 - s_{2l})\mu - x} N\Gamma(l, 0, \dots, 0) + \frac{1}{\lambda\Lambda_N + \mu - x} \sum_{\substack{|\underline{\alpha}|=l \\ \underline{\alpha} \in \Xi}} \Gamma(\underline{\alpha}) \frac{l!}{\alpha_1! \dots \alpha_N!} \\
&\leq \frac{1}{\lambda\Lambda_N + (1 - s_{2l})\mu - x} N\Gamma(l, 0, \dots, 0) + \frac{1}{\lambda\Lambda_N + \mu - x} [1 - N\Gamma(l, 0, \dots, 0)] . \quad (\text{using eq. (29)})
\end{aligned}$$

It is easily seen that the equation

$$\frac{1}{\lambda\Lambda_N} = \frac{1}{\lambda\Lambda_N + (1 - s_{2l})\mu - x} N\Gamma(l, 0, \dots, 0) + \frac{1}{\lambda\Lambda_N + \mu - x} [1 - N\Gamma(l, 0, \dots, 0)] \quad (33)$$

and (27) are equivalent and hence the smallest eigenvalue $a_l \geq x_l$.

Note that necessarily $x_l < \lambda\Lambda_N + \mu(1 - s_{2l})$. Thus, if $K = 0$, $a_l = \lambda\Lambda_N + \mu\sigma_{2\underline{\alpha}}$ for some $\underline{\alpha}$ and hence $a_l \geq \lambda\Lambda_N + \mu(1 - s_{2l}) > x_l$, which proves the theorem. \square

Proof of Theorem 2.2.2. Since symmetric functions are preserved under \mathcal{L} , the space of symmetric Hermite polynomials in L_4 with orthonormal basis

$$\left\{ \sqrt{\frac{2}{N(N-1)}} \sum_{i \neq j} H_2(v_i) H_2(v_j), \sqrt{\frac{2}{3N}} \sum H_4(v_i) \right\}$$

gives rise to two eigenfunctions. The action of $\mu\mathcal{L}_T + \lambda\mathcal{L}_K$ on this space is represented by the following matrix

$$\begin{pmatrix} \mu + \frac{3\lambda}{2(N-1)} & \frac{-\sqrt{3}\lambda}{2\sqrt{N-1}} \\ \frac{-\sqrt{3}\lambda}{2\sqrt{N-1}} & \frac{5\mu}{8} + \frac{\lambda}{2} \end{pmatrix} \quad (34)$$

whose characteristic equation is (17) and smallest eigenvalue is thus a_2 . Hence, we immediately have $\Delta_N^{(2)} \leq a_2$.

To see the opposite inequality recall that x_l is the smaller of the two solutions of the equation (33). Since for $l \geq 2$, $s_{2l} \leq s_4 = \frac{3}{8}$ and $\Gamma(l, 0, \dots, 0) \leq \Gamma(2, 0, \dots, 0) = \frac{3}{N(N+2)}$ we get from (33)

$$\Delta_N^{(2)} \geq a_2 .$$

\square

The eigenfunction corresponding to the “second” gap $\Delta_N^{(2)}$ is given by $\sum_{|\underline{\alpha}|=2} c_{\underline{\alpha}} H_{2\underline{\alpha}} \in L_4$, where $c_{\underline{\alpha}}$ are symmetric under exchange of indices (see eq. (31)). In fact, eq. (31) characterizes the symmetric eigenfunctions of $\mu\mathcal{L}_T + \lambda\mathcal{L}_R$ when $K \neq 0$ and the non-symmetric eigenfunctions when $K = 0$. This means that solutions x of eq. (32) correspond to symmetric eigenfunctions alone. Hence, the unique eigenfunction (by extension, also that of $\mu\mathcal{L}_T + \lambda\mathcal{L}_K$) corresponding to a_2 is symmetric, which is the physically interesting case.

We eventually do get the optimal bound a_2 due to the following reason: The space of symmetric functions in L_4 is spanned by the set $\{\sum_{i \neq j} H_2(v_i)H_2(v_j), \sum H_4(v_i)\}$, which can also be spanned by two functions, one of which is radial (of degree 4) and the other perpendicular to the radial one. The latter gives the gap Λ_N for \mathcal{L}_K . Hence, the action of \mathcal{L}_K and \mathcal{L}_R on the space of symmetric functions in L_4 is precisely the same.

In the limit $N \rightarrow \infty$, the off-diagonal elements of the matrix (34) vanish. Therefore, in this limit, the eigenvalues x_2^\pm tend to $\frac{\lambda}{2} + \frac{5}{8}\mu$ and μ (see (18)), which corresponds to the simultaneous diagonalization of operators \mathcal{L}_T and \mathcal{L}_K .

2.3.2 Model II

The master equation (12) in this case, for the evolution of the perturbation h , becomes

$$\frac{\partial h}{\partial t} = -(\lambda\mathcal{L}_K + \mu\mathcal{L}_P)h, \quad (35)$$

where $\mathcal{L}_K = \mathcal{G}_K$ is the Kac part as defined in eq. (11), and $\mathcal{L}_P := \sum_{j=1}^m (I - P_j)$, where

$$P_j h := \int dw_j g(w_j) h(v_1, \dots, w_j, \dots, v_N).$$

It is easy to see that \mathcal{L}_K and \mathcal{L}_P are self-adjoint on \mathcal{X}_N and each P_j is a projection onto the space of functions in \mathcal{X}_N independent of v_j . Defining the evolution operator $\mathcal{L}_{N,m} := \lambda\mathcal{L}_K + \mu\mathcal{L}_P$, we describe some of its properties.

Lemma 2.3.9.

- $\mathcal{L}_{N,m} \geq 0$ on \mathcal{X}_N .
- $\mathcal{L}_{N,m}h = 0 \Leftrightarrow h = 0$.

Proof. We know from [21] and 2.3.1 that $(I - Q) \geq 0$ and $(I - Q)h = 0 \Leftrightarrow h$ is radial. Each $(I - P_k)$ is a projection with kernel precisely the subspace of functions in \mathcal{X}_N that are independent of v_k . The only function in \mathcal{X}_N that belongs to the kernel of $\sum_{k=1}^m (I - P_k)$ and is also radial is 0. Hence, the Lemma is proved, and we can infer from it that $\gamma(\mathbf{v})$ is the unique equilibrium of (12). \square

Lemma 2.3.9 implies that initial states in \mathcal{X}_N decay to equilibrium at an exponential rate $\delta_{N,m}$, the spectral gap of $\mathcal{L}_{N,m}$ defined in (19).

First, the observation that $\mathcal{L}_{2,1}$ is simply a linear combination of two projections ($Q \equiv R_{12}$ is an orthogonal projection onto radial functions in \mathbb{R}^2) lets us compute the whole spectrum in the two-particle case. This is done in Appendix A. We see that the spectral gap is the lower root of the quadratic $x^2 - (2\lambda + \mu)x + \lambda\mu$:

$$\delta_{2,1} := \frac{(2\lambda + \mu) - \sqrt{4\lambda^2 + \mu^2}}{2} \quad (36)$$

with gap eigenfunction

$$\frac{2\lambda}{2\lambda + \mu - \delta_{2,1}} H_2(v_1) + \frac{2\lambda}{2\lambda - \delta_{2,1}} H_2(v_2),$$

where H_2 is the monic Hermite polynomial (with weight γ) of degree 2.

For general N, m , the asymptotic behavior is given by Theorem 2.2.3, which we now prove.

Proof of Theorem 2.2.3. The proof is based on an inductive argument that follows in essence the one in [7] in which the spectral gap of the Kac model is computed exactly.

We first prove the following claim for $1 \leq m < N$:

$$\delta_{N,m} \geq \frac{N-m-1}{N-1} \delta_{N-1,m} + \frac{m}{N-1} \delta_{N-1,m-1} . \quad (37)$$

We let $\mathcal{L}_{N,m}^{(k)}$ be the evolution operator $\mathcal{L}_{N,m}$ with the k^{th} particle removed:

$$\mathcal{L}_{N,m}^{(k)} = \frac{(N-1)\lambda}{\binom{N-1}{2}} \sum_{\substack{i < j \\ i, j \neq k}}^N (I - R_{ij}) + \mu \sum_{\substack{l=1 \\ l \neq k}}^m (I - P_l).$$

Remark 2.3.10. $\mathcal{L}_{N,m}^{(k)}$ is also self-adjoint in $L^2(\mathbb{R}^N, \gamma)$, and will have m or $m-1$ strong thermostats in it, depending on whether $k > m$ or $k \leq m$, respectively. Also, the coefficient of the Kac term corresponds to collisions among $N-1$ particles.

Next we show that

$$\mathcal{L}_{N,m} = \frac{1}{N-1} \sum_{k=1}^N \mathcal{L}_{N,m}^{(k)}. \quad (38)$$

This follows, since

$$\begin{aligned} \sum_{k=1}^N \mathcal{L}_{N,m}^{(k)} &= \sum_{k=1}^N \left(\frac{2\lambda}{N-2} \sum_{\substack{i < j, \\ i, j \neq k}}^N (I - R_{ij}) + \mu \sum_{\substack{l=1 \\ l \neq k}}^m (I - P_l) \right) \\ &= 2\lambda \sum_{i < j}^N (I - R_{ij}) + (N-1)\mu \sum_{l=1}^m (I - P_l) \\ &= (N-1)\mathcal{L}_{N,m}. \end{aligned}$$

Then

$$\langle h, \mathcal{L}_{N,m}[h] \rangle = \frac{1}{N-1} \sum_{k=1}^N \langle h, \mathcal{L}_{N,m}^{(k)}[h] \rangle . \quad (39)$$

At this point, we want to introduce the gaps $\delta_{N-1,m}$ and $\delta_{N-1,m-1}$ for $N-1$ particles into the right hand side; for this, we will need the functions to be orthogonal to 1 in the space $L^2(\mathbb{R}^{N-1}, \gamma(\hat{v}_k))$, where $\gamma(\hat{v}_k)$ is the Gaussian γ with the variable v_k missing. To this end, we define the projections

$$\pi_k[h] := \int h \gamma(\hat{v}_k) dv_1 \dots dv_{k-1} dv_{k+1} \dots dv_N$$

and write, for each k , $\langle h, \mathcal{L}_{N,m}^{(k)}[h] \rangle = \langle (h - \pi_k h), \mathcal{L}_{N,m}^{(k)}(h - \pi_k h) \rangle$. This holds because the range of the projection π_k is exactly the kernel of $\mathcal{L}_{N,m}^{(k)}$, and the operator $\mathcal{L}_{N,m}^{(k)}$ is self-adjoint. Thus, from (39),

$$\delta_{N,m} = \frac{1}{N-1} \inf \sum_{k=1}^N \langle (h - \pi_k h), \mathcal{L}_{N,m}^{(k)}(h - \pi_k h) \rangle,$$

where the infimum is over $h \in \mathcal{X}_N$, $\|h\| = 1$ as per the definition of the spectral gap. Since $(h - \pi_k h)$ is orthogonal to the constant function 1 in $L^2(\mathbb{R}^{N-1}, \gamma(\hat{v}_k))$ by construction, we use the definition of the spectral gap to write

$$\begin{aligned} \delta_{N,m} &\geq \frac{1}{N-1} \inf \left(\sum_{k=m+1}^N \delta_{N-1,m} (\|h - \pi_k h\|^2) + \sum_{k=1}^m \delta_{N-1,m-1} (\|h - \pi_k h\|^2) \right) \\ &\hspace{25em} \text{(by Remark 2.3.10)} \\ &= \frac{1}{N-1} \inf \left(\delta_{N-1,m} \sum_{k=m+1}^N (\|h\|^2 - \|\pi_k h\|^2) + \delta_{N-1,m-1} \sum_{k=1}^m (\|h\|^2 - \|\pi_k h\|^2) \right) \\ &\geq \frac{N-m}{N-1} \delta_{N-1,m} + \frac{m}{N-1} \delta_{N-1,m-1} - \frac{1}{N-1} \max\{\delta_{N-1,m}, \delta_{N-1,m-1}\} \sup \sum_{k=1}^N \|\pi_k h\|^2, \end{aligned}$$

where we have used symmetry among $1, \dots, m$ and $m+1, \dots, N$ and the fact that the infimum is over functions with norm 1.

First, we note that $\delta_{N-1,m} \geq \delta_{N-1,m-1}$ since $(I - P_m) \geq 0$. Next, $\sup\{\sum_{k=1}^N \|\pi_k h\|^2, h \in \mathcal{X}_N\}$ equals $\sup_{\mathcal{X}_N} \langle h, \sum_{k=1}^N \pi_k h \rangle$. Since $\{\pi_k\}_1^N$ is a collection of commuting projection operators, $\sum_{k=1}^N \pi_k$ is a projection and the supremum is 1.

We then get

$$\delta_{N,m} \geq \frac{N-m}{N-1} \delta_{N-1,m} + \frac{m}{N-1} \delta_{N-1,m-1} - \frac{1}{N-1} \delta_{N-1,m},$$

which implies claim (37).

We now prove the first inequality in Theorem 2.2.3. The region of interest is $\{(N, m) : 1 \leq m \leq N-1\}$. We will use induction on $N \geq 2$.

- The base case $N = 2, m = 1$ is the trivial statement $\delta_{2,1} \geq \delta_{2,1}$.
- Now suppose

$$\delta_{N,m} \geq \delta_{2,1} \frac{m}{N-1} \quad (40)$$

for all m such that $1 \leq m \leq N-1$. To show that $\delta_{N+1,m} \geq \delta_{2,1} \frac{m}{N}$ for all m such that $1 \leq m \leq N$, consider the following two cases:

- $m = 1$: We need to show that $\delta_{N+1,1} \geq \frac{\delta_{2,1}}{N}$. From (37), we deduce that

$$\delta_{N+1,1} \geq \frac{N-1}{N} \delta_{N,1} + \frac{1}{N} \delta_{N,0} = \frac{N-1}{N} \delta_{N,1}.$$

In the above, we have $\delta_{N,0} = 0$ because when none of the particles are thermostated, the ground-state is degenerate (any radial function in \mathbb{R}^N is an equilibrium for the Kac part). Applying (40) with $m = 1$ then completes the proof of this case.

- $1 < m \leq N$:

$$\begin{aligned} \delta_{N+1,m} &\geq \frac{N-m}{N} \left(\frac{m \delta_{2,1}}{N-1} \right) + \frac{m}{N} \left(\frac{(m-1) \delta_{2,1}}{N-1} \right) \quad (\text{using (37) and (40)}) \\ &= \delta_{2,1} \frac{m}{N(N-1)} (N-m+m-1) = \delta_{2,1} \frac{m}{N}. \end{aligned}$$

This proves the first inequality in (20). We prove second inequality in (20), by finding an upper bound proportional to $\frac{m}{N-1}$, for $\delta_{N,m}$. This can be done by finding a (possibly crude) upper bound on the eigenvalues of $\mathcal{L}_{N,m}$ on the space of second

degree Hermite polynomials with weight γ . This space is invariant under $\mathcal{L}_{N,m}$, and the action of $\mathcal{L}_{N,m}$ on it with basis $\{\sum_{k=m+1}^N H_2(v_k), \sum_{k=1}^m H_2(v_k)\}$ can be described by the following matrix. We use the identities $R_{ij}H_2(v_i) = (H_2(v_i) + H_2(v_j))/2$ and $R_{ij}H_2(v_k) = H_2(v_k)$ for $i, j \neq k$ in obtaining the entries.

$$\begin{pmatrix} \frac{\lambda m}{N-1} & \frac{-\lambda m}{N-1} \\ -\frac{\lambda(N-m)}{N-1} & \frac{\lambda(N-m)}{N-1} + \mu \end{pmatrix}$$

Its smallest eigenvalue is $\frac{1}{2}(\mu + \frac{N\lambda}{N-1})(1 - \sqrt{1 - \frac{4m\lambda\mu}{N-1} \frac{1}{(\mu + \frac{N\lambda}{N-1})^2}})$. Hence, by definition of the gap,

$$\delta_{N,m} \leq \frac{1}{2}(\mu + \frac{N\lambda}{N-1})(1 - \sqrt{1 - \frac{m}{N-1} \frac{4\lambda\mu}{(\mu + \frac{N\lambda}{N-1})^2}}).$$

For N large enough, we can write

$$\delta_{N,m} \leq \frac{1}{2} \left(\mu + \frac{N\lambda}{N-1} \right) \frac{m}{N-1} \frac{4\lambda\mu}{(\mu + \frac{N\lambda}{N-1})^2}$$

or

$$\delta_{N,m} \leq \frac{m}{N-1} \frac{2\lambda\mu}{\mu + \lambda}.$$

□

Thus, as we are close to equilibrium, $h \rightarrow 0$ in $L^2(\mathbb{R}^N, \gamma)$ at an exponential rate $\delta_{N,m}$, which for large N , is proportional to the fraction of thermostated particles.

Remark. The matrix in the previous proof is related to the evolution of kinetic energy of the system, see eq. (15), as the behavior of the kinetic energy is indicative of the action of the operator $\mathcal{L}_{N,m}$ on polynomials of the form v_j^2 . Moreover, for $N = 2, m = 1$, we show in Appendix A that the gap eigenfunction - the slowest rate of decay in the space $L^2(\mathbb{R}^2, \gamma)$ - is a second degree polynomial. One may thus wonder if the gap eigenfunction is a second degree polynomial for other values of N too. However, currently we only have asymptotic bounds on $\delta_{N,m}$.

2.4 Approach to Equilibrium in Entropy

In this section, we study the behavior of the relative entropy functional

$$S(f|\gamma) := \int f \log \frac{f}{\gamma} d\mathbf{v} \quad (41)$$

under the two evolutions (11) and (12). We have

$$\begin{aligned} \int f \log \frac{f}{\gamma} d\mathbf{v} &= \int \left(\frac{f}{\gamma}\right) \log\left(\frac{f}{\gamma}\right) \gamma d\mathbf{v} \\ &\geq \left(\int \frac{f}{\gamma} \gamma d\mathbf{v}\right) \log\left(\int \frac{f}{\gamma} \gamma d\mathbf{v}\right) \end{aligned}$$

by Jensen's inequality, and this implies that for any probability distribution f on \mathbb{R}^N , $S(f|\gamma) \geq 0$. Moreover, $S(f|\gamma) = 0 \Leftrightarrow f = \gamma$. Hence, the tracking the decrease in $S(f|\gamma)$ as f evolves is a good measure of equilibration.

We are interested in quantifying the decay in S , for which one approach has been to study the entropy production $-\frac{dS}{dt}$ along the evolution, and in particular, computing $\sup_f -\frac{1}{S} \frac{dS(f(.,t)|\gamma)}{dt}$, as $\frac{dS}{dt}$ is linear in the generator of the evolution (see eq. (42)). However, our approach is to look at the term $S(f(.,t)|\gamma)$ and utilize its convexity. This lends itself to the more general case where the entropy decay is sensed only in higher order terms, and not in the linear term.

We use the representation $f = \gamma h$, where $h(\mathbf{v}) = 1$ indicates equilibrium. This is slightly different from the ground-state transformation (16). Since f is a probability distribution, we restrict to $h \geq 0$, with $\int h \gamma d\mathbf{v} = 1$. The relative entropy then becomes $\int h \log h \gamma d\mathbf{v}$, which we denote by $S(h)$ (overloading the notation) for the remainder of this section.

The main results in this section are the proofs of Theorems 2.2.4 and 2.2.5.

2.4.1 Model I

The evolution of h is given by (23), which we restate below.

$$\frac{\partial h}{\partial t} = N\lambda(Q - I)h + \mu \sum_{j=1}^N (T_j - I)h .$$

To begin, we show that $S(h)$ decreases under this evolution.

Now,

$$\frac{dS}{dt} = \int \frac{\partial h}{\partial t} \log h \gamma d\mathbf{v} + \int \frac{h}{h} \frac{\partial h}{\partial t} \gamma d\mathbf{v} = \int \frac{\partial h}{\partial t} \log h \gamma d\mathbf{v} , \quad (42)$$

where the second term vanishes because the normalization $\int h \gamma d\mathbf{v} = 1$ is preserved by the evolution. Relative entropy decreases along the Kac flow ([21]), and since

$$\begin{aligned} & \sum_j \int (T_j - I) h \log h \gamma d\mathbf{v} \\ &= \sum_j \int T_j h \log h \gamma d\mathbf{v} - \frac{N}{2} S(h) \\ &= \frac{1}{2} \sum_j \left(\int T_j h \log h \gamma d\mathbf{v} + \int h T_j (\log h) \gamma d\mathbf{v} \right) - \frac{N}{2} S(h) \\ &\leq \frac{1}{2} \sum_j \left(\int T_j h \log h \gamma d\mathbf{v} + \int h \log(T_j h) \gamma d\mathbf{v} \right) - \frac{N}{2} S(h) \\ &\quad \text{(by concavity of log and averaging property of } T_j) \\ &= \frac{1}{2} \sum_j \left(\int T_j h \log h \gamma d\mathbf{v} + \int h \log(T_j h) \gamma d\mathbf{v} - \int h \log h \gamma d\mathbf{v} - \int (T_j h) \log(T_j h) \gamma d\mathbf{v} \right) \\ &\quad + \frac{1}{2} \sum_j \left(\int h \log h \gamma d\mathbf{v} + \int (T_j h) \log(T_j h) \gamma d\mathbf{v} \right) - \frac{N}{2} S(h) \\ &\leq \frac{1}{2} \sum_j \int (h - (T_j h)) (\log(T_j h) - \log(h)) \gamma d\mathbf{v} + \frac{N}{2} S(h) - \frac{N}{2} S(h) \\ &\quad \text{(by convexity of } S \text{ as shown in Lemma 2.4.1)} \\ &\leq 0 , \end{aligned}$$

we have that $\frac{dS}{dt} \leq 0$ for Model I. Theorem 2.2.4 makes a stronger claim: that if $h(\mathbf{v}, t)$ is a solution of eq. (23), then $S(h(\cdot, t))$ decays to 0 exponentially as $t \rightarrow \infty$ via

$$S(h(\cdot, t)) \leq e^{-\rho t} S(h(\cdot, 0)) , \quad (43)$$

where $\rho = \frac{\mu}{2}$.

The idea is to write $h(\mathbf{v}, t)$ explicitly as $e^{-(N\lambda(Q-I) + \mu \sum_{j=1}^N (T_j - I))t} h(\mathbf{v}, 0)$, utilize convexity of $S(h)$ to expand the exponential, and prove entropy decay in each term of the expansion. The upcoming lemmas provide the necessary ingredients.

Lemma 2.4.1. *Let $a, b \in \mathbb{R}$ such that $a + b = 1$. Let h_1, h_2 be such that $\int \gamma h_1 = \int \gamma h_2 = 1$, and $h_1, h_2 \geq 0$. Then*

$$S(ah_1 + bh_2) \leq aS(h_1) + bS(h_2) .$$

Proof. This follows immediately from Jensen's inequality and the fact that $S(h)$ is an integral of a convex function. \square

Lemma 2.4.2. *Given $h \in L^1(\mathbb{R}^N, \gamma(\mathbf{v})d\mathbf{v})$ such that $\int h\gamma d\mathbf{v} = 1$ and $h \geq 0$,*

$$S(Qh) \leq S(h) .$$

Proof. The entropy is non-expansive under the rotations $R_{ij}h = \int h(\dots, v_i^*, \dots, v_j^*, \dots) d\theta$ as shown below:

$$\begin{aligned} S(R_{ij}h) &= \int (R_{ij}h) \log(R_{ij}h) \gamma d\mathbf{v} \\ &\leq \int R_{ij} (h \log h) \gamma d\mathbf{v} , \end{aligned}$$

by Jensen's inequality. The claim follows as γ is invariant under rotations. Now by convexity of entropy (Lemma 2.4.1),

$$S(Qh) \leq \frac{1}{\binom{N}{2}} \sum_{i < j} S(R_{ij}h) \leq \frac{1}{\binom{N}{2}} S(h) = S(h) .$$

\square

The next proposition shows that in the case of one-particle ($N = 1$), the entropy is a strict contraction under the action of the thermostat, with coefficient $\frac{1}{2}$.

Proposition 2.4.3. *Given $h \in L^1(\mathbb{R}, g(v)dv)$, with $\int h(v)g(v)dv = 1$, $h \geq 0$ and*
 $Th := \int dw g(w) \int_0^{2\pi} d\theta h(v \cos \theta + w \sin \theta),$

$$S(Th) \leq \frac{1}{2} S(h) .$$

To prove this, we will be invoking the following well-known property of the Ornstein-Uhlenbeck process, see [28, 16, 1, 17, 29].

Theorem 2.4.4. *Let P_s be the semigroup generated by the 1-dimensional Ornstein-Uhlenbeck process, that is, $U_s = P_s[U_0]$ is the solution of the Fokker-Planck equation*

$$\frac{\partial U_s(v)}{\partial s} = U_s''(v) - vU_s'(v)$$

with initial condition U_0 . For every density h on $L^1(\mathbb{R}, g(v)dv)$ we have

$$\int g(v)dv P_s[h](v) \log(P_s[h](v)) \leq e^{-2s} \int g(v)dv h(v) \log h(v) .$$

Remark. The semigroup, which can be represented explicitly as

$$P_s[h](v) = \int dw g(w) h(e^{-s}v + \sqrt{1 - e^{-2s}}w) , \quad (44)$$

is self-adjoint in $L^2(\mathbb{R}, g(v)dv)$.

We are now ready to prove Proposition 2.4.3.

Proof of Proposition 2.4.3. To connect the Ornstein-Uhlenbeck process P_s with the operator T we set

$$\bar{T}[h](v) := \int dw g(w) \int_0^{\pi/2} d\theta h(v \cos \theta + w \sin \theta) = \frac{2}{\pi} \int_0^\infty ds \frac{e^{-s}}{\sqrt{1 - e^{-2s}}} P_s[h](v) ,$$

where we use eq. (44) and the change of variables $\cos(\theta) = e^{-s}$. It follows that

$$\begin{aligned} \int dv g(v) \bar{T}[h] \log \bar{T}[h] &= \int dv g(v) \left(\frac{2}{\pi} \int_0^\infty ds \frac{e^{-s}}{\sqrt{1 - e^{-2s}}} P_s[h] \right) \log \left(\frac{2}{\pi} \int_0^\infty ds' \frac{e^{-s'}}{\sqrt{1 - e^{-2s'}}} P_{s'}[h] \right) \\ &\leq \int dv g(v) \left(\frac{2}{\pi} \int_0^\infty ds \frac{e^{-s}}{\sqrt{1 - e^{-2s}}} P_s[h] \log P_s[h] \right) \\ &\quad \text{(using convexity of } x \log x \text{)} \\ &\leq \frac{2}{\pi} \int_0^\infty ds \frac{e^{-s}}{\sqrt{1 - e^{-2s}}} e^{-2s} \int dv g(v) h \log h \quad \text{(using Theorem 2.4.4)} \\ &= \frac{1}{2} \int dv g(v) h \log h . \end{aligned}$$

The next step is to prove the corresponding result for the operator T . Let $h = h_e + h_o$ where h_e is even, i.e. $h_e(v) = h_e(-v)$, and h_o is odd, i.e. $h_o(-v) = -h_o(v)$. Observe that $T[h]$ is even, $T[h_o] = 0$ and $\bar{T}[h_e] = T[h_e]$. While the first two identities follow directly from the definitions, the last one also uses the fact that $\int dw g(w) \int_{\frac{\pi}{2}}^{\pi} d\theta h_e(v \cos \theta + w \sin \theta) = \int dw g(w) \int_0^{\frac{\pi}{2}} d\theta h_e(-v \cos \theta - w \sin \theta)$ under the change of variables $\theta \rightarrow \pi - \theta$ and $w \rightarrow -w$. Thus,

$$\begin{aligned} \int dv g(v) T[h](v) \log T[h](v) &= \int dv g(v) T[h_e](v) \log T[h_e](v) \\ &= \int dv g(v) \bar{T}[h_e](v) \log \bar{T}[h_e](v) \\ &\leq \frac{1}{2} \int dv g(v) h_e(v) \log h_e(v) \\ &\leq \frac{1}{2} \int dv g(v) h(v) \log h(v) , \end{aligned}$$

where, in the last inequality, we have used that $h_e(v) = (h(v) + h(-v))/2$ and Jensen's inequality. \square

Remark 2.4.5. The above Proposition is a short step from proving Theorem 2.2.4 for the $N = 1$ case. To be precise,

$$\begin{aligned} \int dv g(v) e^{(T-I)t} h \log(e^{(T-I)t} h) &\leq e^{-t} \sum_{k=0}^{\infty} \frac{t^k}{k!} \int dv g(v) T^k[h](v) \log T^k[h](v) \quad (\text{by convexity}) \\ &\leq e^{-t} \sum_{k=0}^{\infty} \left(\frac{t}{2}\right)^k \frac{1}{k!} \int dv g(v) h(v) \log h(v) \quad (\text{by Proposition 2.4.3}) \\ &= e^{-\frac{t}{2}} \int dv g(v) h(v) \log h(v) . \end{aligned}$$

The following two lemmas will help extend the result to $N > 1$.

Lemma 2.4.6. *Let $h(\mathbf{v})$ satisfy the assumptions in Lemma 2.4.2. Then*

$$\sum_{i=1}^N S(P_i h) \leq (N-1) S(h) .$$

The above Lemma is a version of Han's inequality [18] adapted to our situation. We first prove the following useful Proposition.

Proposition 2.4.7. Let $f(\mathbf{v})$ be a probability density on \mathbb{R}^N and let its marginal over the j^{th} variable be denoted by $f_j(\hat{\mathbf{v}}_j) = \int f(\mathbf{v}) dv_j$, where $\hat{\mathbf{v}}_j = (v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_N)$.

Then we have

$$\sum_{j=1}^N \int f_j \log f_j d\hat{\mathbf{v}}_j \leq (N-1) \int f \log f d\mathbf{v} .$$

Proof. We first observe that from the Loomis-Whitney inequality [24], that is

$$\int_{\mathbb{R}^N} F_1(\hat{\mathbf{v}}_1) \dots F_N(\hat{\mathbf{v}}_N) \leq \|F_1\|_{L^{N-1}} \dots \|F_N\|_{L^{N-1}}$$

for $F_j \in L^{N-1}(\mathbb{R}^{N-1})$, it follows that

$$Z := \int \prod_{j=1}^N f_j^{\frac{1}{N-1}} d\mathbf{v} \leq 1 . \quad (45)$$

Thus we have

$$\begin{aligned} \int f \log \left[\frac{f}{\prod f_j^{\frac{1}{N-1}}} \right] d\mathbf{v} &= Z \int \frac{f}{\prod f_j^{\frac{1}{N-1}}} \log \left[\frac{f}{\prod f_j^{\frac{1}{N-1}}} \right] \frac{\prod f_j^{\frac{1}{N-1}}}{Z} d\mathbf{v} \\ &\geq Z \left[\int \frac{f}{Z} d\mathbf{v} \right] \log \left[\int \frac{f}{Z} d\mathbf{v} \right] = -\log Z , \end{aligned}$$

where we have used Jensen's inequality and the convexity of $x \log(x)$. The Lemma follows easily from the above inequality and (45). \square

Proof of Lemma 2.4.6. Moving to the “ f ” representation for the moment ($f = \gamma h$), and writing

$$P_i h = \frac{1}{\gamma(\hat{v}_i)} \int f(\dots v_i \dots) dv_i ,$$

we have

$$\begin{aligned} \sum_{i=1}^N S(P_i h) &= \sum \int f_i(\hat{v}_i) \log \left(\frac{f_i(\hat{v}_i)}{\gamma(\hat{v}_i)} \right) d\hat{v}_i \\ &= \sum \int f_i \log f_i d\hat{v}_i - \sum \int f_i \log \gamma(\hat{v}_i) d\hat{v}_i \\ &\leq (N-1) \int f \log f d\mathbf{v} - (N-1) \int f \log \gamma d\mathbf{v} , \end{aligned}$$

where the last step follows from Proposition 2.4.7 and the following computation

$$\sum \int f_i \log \gamma(\hat{v}_i) d\hat{v}_i = \sum \int f \log \gamma(\hat{v}_i) d\mathbf{v} = \int f \log \left(\prod \gamma(\hat{v}_i) \right) d\mathbf{v} .$$

Switching back to the “ h ” representation, we have the result:

$$\sum_{i=1}^N S(P_i h) \leq (N-1) \int f \log \frac{f}{\gamma} d\mathbf{v} = (N-1) S(h) .$$

□

Lemma 2.4.8. *Let $h(\mathbf{v})$ satisfy the assumptions in Lemma 2.4.2. Then*

$$S \left(\frac{1}{N} \sum_{i=1}^N T_i h \right) \leq \left(\frac{N - \frac{1}{2}}{N} \right) S(h) .$$

Proof. First, we have

$$\begin{aligned} S(T_i h) &= \int (T_i h) \log(T_i h) \gamma d\mathbf{v} \\ &= \int \left(\frac{T_i h}{P_i h} \right) \log \left(\frac{T_i h}{P_i h} \right) (P_i h) \gamma d\mathbf{v} + \int T_i h \log(P_i h) \gamma d\mathbf{v} \\ &= \int T_i \left(\frac{h}{P_i h} \right) \log \left(T_i \left(\frac{h}{P_i h} \right) \right) (P_i h) \gamma d\mathbf{v} + \int T_i h \log(P_i h) \gamma d\mathbf{v} \\ &\quad (\text{since } P_i h \text{ is independent of } v_i) \\ &\leq \frac{1}{2} \int \frac{h}{P_i h} \log \left(\frac{h}{P_i h} \right) (P_i h) \gamma d\mathbf{v} + \int T_i h \log(P_i h) \gamma d\mathbf{v} , \end{aligned}$$

where the last step follows by applying Proposition 2.4.3 to $\frac{h}{P_i h}$ as a function of v_i alone. It satisfies the assumptions since $\int \frac{h}{P_i h} g(v_i) dv_i = \frac{P_i h}{P_i h} = 1$. Thus, so far, we have

$$\begin{aligned} S(T_i h) &\leq \frac{1}{2} \int h \log \left(\frac{h}{P_i h} \right) \gamma d\mathbf{v} + \int h \log(P_i h) \gamma d\mathbf{v} \\ &= \frac{1}{2} S(h) + \frac{1}{2} S(P_i h) . \end{aligned}$$

The $T_i h$ in the second term could be replaced by h due to self-adjointness of T_i in $L^2(\mathbb{R}^N, \gamma d\mathbf{v})$ and the fact that $\log P_i h$ is independent of v_i . Now, we start with the

left-hand side in the statement of the Lemma, and apply convexity of $S(h)$ (from Lemma 2.4.1):

$$\begin{aligned} S\left(\frac{1}{N} \sum_{i=1}^N S(T_i h)\right) &\leq \frac{1}{N} \sum_{i=1}^N S(T_i h) \\ &\leq \frac{1}{2} S(h) + \frac{1}{2N} \sum_{i=1}^N S(P_i h), \end{aligned}$$

where the last step follows from the previous computation. Finally, applying Lemma 2.4.6 to the second term yields the proof. \square

Remark. The previous Lemma 2.4.6 implies that $S\left(\frac{1}{N} \sum_{i=1}^N P_i h\right) \leq \left(\frac{N-1}{N}\right) S(h)$. Notice that this is slightly stronger than the result of the above Lemma. This is not surprising, as P_i can be interpreted as a “stronger” counterpart of T_i .

We are now ready to prove the main theorem of this section.

Proof of Theorem 2.2.4. Writing the exponential of the generator of the evolution as a convex combination,

$$e^{-(N\lambda(Q-I) + \mu \sum_{j=1}^N (T_j - I))t} = \frac{N^k (\mu + \lambda)^k t^k}{k!} \left(\frac{\mu}{\mu + \lambda} \frac{1}{N} \sum_{i=1}^N T_i + \frac{\lambda}{\mu + \lambda} Q \right)^k.$$

Then, Lemma 2.4.1 implies that

$$S(h(\cdot, t)) \leq e^{-Nt(\mu + \lambda)} \sum_{k=0}^{\infty} \frac{N^k (\mu + \lambda)^k t^k}{k!} S \left[\left(\frac{\mu}{\mu + \lambda} \frac{1}{N} \sum_{i=1}^N T_i + \frac{\lambda}{\mu + \lambda} Q \right)^k h(\cdot, 0) \right]. \quad (46)$$

We digress to analyze the term with the square bracket, and will return to (46) after.

$$\begin{aligned} S \left[\left(\frac{\mu}{\mu + \lambda} \frac{1}{N} \sum_{i=1}^N T_i + \frac{\lambda}{\mu + \lambda} Q \right) h \right] &\leq \frac{\mu}{\mu + \lambda} S \left(\frac{1}{N} \sum_{i=1}^N T_i h \right) + \frac{\lambda}{\mu + \lambda} S(Qh) \\ &\leq \frac{\mu}{\mu + \lambda} \frac{N - \frac{1}{2}}{N} S(h) + \frac{\lambda}{\mu + \lambda} S(h), \end{aligned}$$

where the last step is inferred from Lemmas 2.4.8 and 2.4.2. Having shown that

$$S \left[\left(\frac{\mu}{\mu + \lambda} \frac{1}{N} \sum_{i=1}^N T_i + \frac{\lambda}{\mu + \lambda} Q \right)^k h(., 0) \right] \leq \left(\frac{\mu}{\mu + \lambda} \frac{N - \frac{1}{2}}{N} + \frac{\lambda}{\mu + \lambda} \right)^k S(h(., 0)) ,$$

substituting the above in eq. (46) yields the result eq. (43). \square

In contrast to the Kac model, the presence of the thermostat guarantees an exponential rate of convergence strictly positive uniformly in N . It is fundamental in the above analysis that the thermostat acts on all particles. The presence of the Kac part gives no contribution to the above estimate of the exponential decay rate (see Lemma 2.4.2).

Comparing the first order terms in the Taylor expansion of the statement of the main Theorem 2.2.4 shows that

$$\frac{dS(f(\mathbf{v}, t)|\gamma)}{dt} \leq -\frac{\mu}{2} S(f(\mathbf{v}, t)|\gamma) . \quad (47)$$

This can be leveraged to yield a lower bound on the spectral gap Δ_N as follows: given a function f of the form

$$f = \gamma(1 + \epsilon h)$$

with $\int h \gamma = 0$ and ϵ small, one can write

$$\epsilon \int \gamma \frac{\partial h}{\partial t} \left(\epsilon h - \frac{\epsilon^2 h^2}{2} \dots \right) \leq -\rho \int \gamma(1 + \epsilon h) \left(\epsilon h - \frac{\epsilon^2 h^2}{2} \dots \right) ,$$

where $\rho = \mu/2$. That is,

$$\int \gamma h \frac{\partial h}{\partial t} \leq -\rho \int \frac{\gamma h^2}{2} .$$

Thus in $L^2(\mathbb{R}^N, \gamma(\mathbf{v}) d\mathbf{v})$ we get

$$\frac{d}{dt} \|h\| \leq -\frac{\rho}{2} \|h\| .$$

Observe that this is very similar to the result one get from Proposition 2.2.1 but $\rho < \mu$. One may wonder whether ρ is the optimal estimate for the decay rate of the

relative entropy. We show that it is so for the case $N = 1$ through the following optimizing sequence similar to that used in [2, 6, 13] (It is easy to generalize the following to the $N > 1$ case in which the Kac term exists). Consider

$$\phi_\delta(v) := (1 - \delta)M_x(v) + \delta M_y(v) ,$$

where $x = \frac{1}{(1-\delta)}$, $y = \frac{1}{\delta}$ and $M_a(v) = \frac{1}{\sqrt{2\pi a}} e^{-v^2/2a}$. The governing evolution equation is

$$\frac{\partial f}{\partial t} = \mu(W - I)f .$$

We claim that

$$\lim_{\delta \rightarrow 0} \frac{1}{S(\phi_\delta)} \frac{dS}{dt}(\phi_\delta) \geq -\frac{\mu}{2} ,$$

thereby showing that (47) is an optimal bound. ϕ_δ is a convex combination of Maxwellians, one of which approaches the distribution of the heat bath $M_1 = g$ and the other corresponds to a very high energy distribution (albeit with a vanishing weight) as $\delta \rightarrow 0$. Notice also that the two functions $(1 - \delta)M_x$ and δM_y contribute equally to the total kinetic energy. These types of functions have been used in [2, 6, 13] as examples of distributions that are away from equilibrium (in the sense of the entropy) and yet have vanishingly low entropy production (in magnitude) with respect to the Kac-Boltzmann equation. Another interesting fact about ϕ_δ is that $S(\phi_\delta|\gamma) \rightarrow \frac{1}{2}$ as $\delta \rightarrow 0$ (this is shown below), but $S(\gamma|\phi_\delta) \rightarrow 0$.

From eq. (47), the entropy production for the above evolution satisfies the bound

$$\frac{dS}{dt} \leq -\frac{\mu}{2} S .$$

In other words, we have that

$$\int W f \log \frac{f}{g} dv \leq \frac{1}{2} \int f \log \frac{f}{g} dv . \quad (48)$$

Here, we show that the sequence ϕ_δ satisfies

$$\lim_{\delta \rightarrow 0} \int W \phi_\delta \log \frac{\phi_\delta}{g} dv \geq \lim_{\delta \rightarrow 0} \frac{1}{2} \int \phi_\delta \log \frac{\phi_\delta}{g} dv \quad (49)$$

while $S(\phi_\delta|g)$ remains bounded away from 0, and hence would work as an optimizing sequence, as claimed. First, we study the behavior of $S(\phi_\delta|g)$ as $\delta \rightarrow 0$. We will occasionally drop the subscript δ .

$$\begin{aligned} S(\phi|g) &= \int \frac{\phi}{g} \log \frac{\phi}{g} g dv \\ &\leq (1-\delta)S(M_x|g) + \delta S(M_y|g) \\ &= (1-\delta)\left(\frac{1}{2} \log(1-\delta) + \frac{1}{2(1-\delta)} - \frac{1}{2}\right) + \delta\left(\frac{1}{2} \log \delta + \frac{1}{2\delta} - \frac{1}{2}\right). \end{aligned}$$

Hence, $\lim_{\delta \rightarrow 0} S(\phi|g) \leq \frac{1}{2}$. However,

$$\begin{aligned} S(\phi|g) &= (1-\delta) \int M_x \log \frac{\phi}{g} + \delta \int M_y \log \frac{\phi}{g} \\ &\geq (1-\delta) \int M_x \log \frac{(1-\delta)M_x}{g} + \delta \int M_y \log \frac{\delta M_y}{g} \quad (\text{by monotonicity of log}) \\ &= (1-\delta) \log(1-\delta) + \delta \log \delta + (1-\delta)S(M_x|g) + \delta S(M_y|g). \end{aligned}$$

Following the previous computation, we get $\lim_{\delta \rightarrow 0} S(\phi|g) \geq \frac{1}{2}$. Hence, $\lim_{\delta \rightarrow 0} S(\phi|g) = \frac{1}{2}$. This is finite and bounded away from zero for δ small enough. We have thus computed the right hand side of (49).

Now, we observe that by (48), $\int W \phi \log \frac{\phi}{g} dv$ is finite for all δ small enough. In the following, we estimate this as $\delta \rightarrow 0$ and prove (49).

$$\begin{aligned} \int W \phi \log \frac{\phi}{g} dv &= (1-\delta) \int W M_x \log \frac{\phi}{g} dv + \delta \int W M_y \log \frac{\phi}{g} dv \\ &\geq (1-\delta) \int W M_x \log \frac{(1-\delta)M_x}{g} dv + \delta \int W M_y \log \frac{\delta M_y}{g} dv \\ &= (1-\delta) \left[\log(1-\delta) + \int W M_x \log \frac{M_x}{g} \right] + \delta \left[\log \delta + \int W M_y \log \frac{M_y}{g} \right] \\ &= (1-\delta) \left[\log(1-\delta) + \frac{1}{2} S(M_x|g) + K_x \right] + \delta \left[\log \delta + \frac{1}{2} S(M_y|g) + K_y \right], \end{aligned}$$

where $K_c := \frac{1}{4}(1 - \frac{1}{2c}) - \frac{1}{4} \log(2c)$ and we have that $\int W M_c \log \frac{M_c}{g} = \frac{1}{2} S(M_c|g) + K_c$.

Thus,

$$\begin{aligned}
\int W \phi \log \frac{\phi}{g} dv &\geq (1 - \delta) \left(\frac{1}{2} \int M_x \log \left(\frac{(1 - \delta) M_x}{g} \right) + K_x + \frac{1}{2} \log(1 - \delta) \right) \\
&\quad + \delta \left(\frac{1}{2} \int M_y \log \left(\frac{\delta M_y}{g} \right) + K_b + \frac{1}{2} \log \delta \right) \\
&= \frac{1}{2} (S((1 - \delta) M_x | g) + S(\delta M_y | g)) \\
&\quad + (1 - \delta) \left(K_x + \frac{1}{2} \log(1 - \delta) \right) + \delta \left(K_y + \frac{1}{2} \log \delta \right) \\
&\geq \frac{1}{2} S(\phi | g) + (1 - \delta) \left(K_x + \frac{1}{2} \log(1 - \delta) \right) + \delta \left(K_y + \frac{1}{2} \log \delta \right) .
\end{aligned}$$

Taking limits, (49) is proved.

Hence, we have shown that (47) is the best bound for the exponential decay rate. This bound does not involve the parameter λ of the Kac collision, though, so one could wonder if the role of the Kac term can be quantified by studying higher derivatives of the entropy.

Remark 2.4.9. All the results in this subsection can be proved in terms of the entropy production, which is the approach used in [3]. However, this same is not so for the next subsection, which deals with Model II.

2.4.2 Model II

The evolution equation obeyed by $h(\mathbf{v}, t)$ is eq. (35), which we restate below:

$$\frac{\partial h}{\partial t} = N\lambda(Q - I)h + \mu \sum_{k=1}^m (P_k - I)h = -\mathcal{L}_{N,m}h .$$

Here,

$$\frac{dS}{dt} = \int \left(N\lambda(Q - I)h + \mu \sum_{k=1}^m (P_k - I)h \right) \log h \gamma d\mathbf{v} .$$

We know (from [21]) that $\int N(Q - I)h \log h \gamma d\mathbf{v} \leq 0$. Also,

$$\begin{aligned} \int P_k h \log h \gamma d\mathbf{v} &= \int P_k h P_k (\log h) \gamma d\mathbf{v} \quad (\text{by self-adjointness of } P_k \text{ as observed in Section 2.3}) \\ &\leq \int (P_k h) \log(P_k h) \gamma d\mathbf{v} \quad (\text{by concavity of } \log \text{ and averaging property of } P_k) \\ &\leq \int h \log h \gamma d\mathbf{v} \quad (\text{by convexity of } x \log x) . \end{aligned}$$

Thus $\frac{dS}{dt} \leq 0$. Theorem 2.2.5 describes the decay of the relative entropy quantitatively, through the equation

$$S(h(., t)) \leq \left(-\frac{\xi_-}{\xi_+ - \xi_-} e^{-\xi_+ t} + \frac{\xi_+}{\xi_+ - \xi_-} e^{-\xi_- t} \right) S(h(., 0)) .$$

Here are a few remarks on this bound. Recall that

$$Z(t) = -\frac{\xi_- e^{-\xi_+ t}}{\xi_+ - \xi_-} + \frac{\xi_+ e^{-\xi_- t}}{\xi_+ - \xi_-}$$

is identically equal to 1 when λ or μ is 0. For $\lambda, \mu > 0$, $\lim_{t \rightarrow \infty} Z(t) = 0$, $Z(t)$ is equal to 1 at $t = 0$ and it is a decreasing function of $t > 0$. The last claim can be seen by computing

$$\frac{dZ}{dt} = \frac{\xi_- \xi_+}{\xi_+ - \xi_-} (e^{-\xi_+ t} - e^{-\xi_- t}) \leq 0 \quad (50)$$

since $\xi_- < \xi_+$. For large t , the dominant term in the bound (22) is $e^{-\xi_- t}$, and for large N , $\xi_- \sim \frac{m\lambda\mu}{(N-1)(N\lambda+\mu)}$. Hence, this bound yields an eventually exponential decay of relative entropy, although the decay rate $\sim \frac{m}{N^2}$ vanishes in the macroscopic limit $N \rightarrow \infty$ even if m is a finite fraction of N .

For the special case $N = 2, m = 1$, observe that $\xi_-(2, 1) = \delta_{2,1}$ is the spectral gap of $2\lambda(I - Q) + \mu(I - P_1)$ (see (36)).

The Theorem is proved as follows: we write $h(\mathbf{v}, t)$ explicitly in terms of the exponential of the generator of the evolution, expand the latter using the Dyson series and use the convexity of the entropy. We exploit the entropic *contraction* of terms of the form $P_j Q$ in the expansion. These steps will yield a non-trivial bound for the entropy at time t in terms of the initial entropy.

The following lemmas build up to the evolution operator $e^{-\mathcal{L}_{N,m}t}$ via the terms in the Dyson expansion in steps. For instance, Lemma 2.4.10 bounds some of the terms obtained by decomposing the Kac operator in the expression $S(P_1 Qh)$. Throughout, we assume that $h \in L^1(\mathbb{R}^N, \gamma)$ and $h \geq 0$.

Lemma 2.4.10. *We have*

$$\sum_{j=2}^N S(P_1 R_{1j} h) \leq \left((N-1) - \frac{1}{2}\right) S(h) .$$

Remark. It is interesting to compare this with Lemma 2.4.8.

Proof. In the following proof, we will apply the continuous version of Han's inequality [18] (this also follows from the Loomis-Whitney inequality [24], as shown in Lemma 2.4.6) for the entropy rewritten to suit our situation:

$$\sum_{j=1}^N S(P_j h) \leq (N-1) S(h) . \quad (51)$$

Note that if h is symmetric in its arguments, this amounts to saying that for each $j = 1, \dots, N$,

$$S(P_j h) \leq \frac{N-1}{N} S(h) . \quad (52)$$

For $j > 1$,

$$\begin{aligned} S(P_1 R_{1j} h) &= \int P_1 R_{1j} h \log(P_1 R_{1j} h) \gamma d\mathbf{v} \\ &= \int P_1 \left(\frac{R_{1j} h}{P_1 P_j h}\right) \log\left(P_1 \left(\frac{R_{1j} h}{P_1 P_j h}\right)\right) P_1 P_j h \gamma d\mathbf{v} + \int P_1 R_{1j} h \log(P_1 P_j h) \gamma d\mathbf{v} , \end{aligned}$$

where we use that $P_1 P_j h$ does not depend on v_1 . Since the argument of the logarithm in the last term is also independent of v_j , we can integrate $P_1 R_{1j} h$ with respect to those variables and use that $\int P_1 R_{1j} h g(v_1) g(v_j) dv_1 dv_j = \int h g(v_1) g(v_j) dv_1 dv_j = P_1 P_j h$ to write:

$$S(P_1 R_{1j} h) = \int P_1 \left(\frac{R_{1j} h}{P_1 P_j h}\right) \log\left(P_1 \left(\frac{R_{1j} h}{P_1 P_j h}\right)\right) P_1 P_j h \gamma d\mathbf{v} + \int P_1 P_j h \log(P_1 P_j h) \gamma d\mathbf{v} .$$

Now, we apply the symmetric version of Han's inequality (52) to $\frac{R_{1j}h}{P_1P_jh}$ as a function of v_1 and v_j to get:

$$\begin{aligned} S(P_1R_{1j}h) &\leq \frac{1}{2} \int \frac{R_{1j}h}{P_1P_jh} \log\left(\frac{R_{1j}h}{P_1P_jh}\right) P_1P_jh \gamma d\mathbf{v} + \int P_1P_jh \log(P_1P_jh) \gamma d\mathbf{v} \\ &= \frac{1}{2} S(R_{1j}h) - \frac{1}{2} \int R_{1j}h \log(P_1P_jh) \gamma d\mathbf{v} + \int P_1P_jh \log(P_1P_jh) \gamma d\mathbf{v} \\ &= \frac{1}{2} S(R_{1j}h) + \frac{1}{2} S(P_1P_jh), \end{aligned}$$

where, to get to the last step, we have used that R_{1j} is self-adjoint and P_1P_j is independent of v_1 and v_j .

Now, summing these terms, and noting that $S(R_{1j}h) \leq S(h)$ by the averaging property of R_{1j} , we get

$$\sum_{j=2}^N S(P_1R_{1j}h) \leq \frac{N-1}{2} S(h) + \frac{1}{2} \sum_{j=2}^N S(P_jP_1h).$$

We invoke Han's inequality (51) on $P_1h \equiv (P_1h)(v_2, \dots, v_N)$, ie. $\sum_{j=2}^N S(P_jP_1h) \leq (N-2)S(P_1h) \leq (N-2)S(h)$ to complete the proof. \square

Lemma 2.4.11.

$$S(e^{\mu(P_1-I)t}Qh) \leq \left(1 - \frac{1 - e^{-\mu t}}{N(N-1)}\right) S(h).$$

Proof.

$$\begin{aligned} S(e^{\mu(P_1-I)t}Qh) &= S(e^{-\mu t}Qh + (1 - e^{-\mu t})P_1Qh) \quad (\text{since } P_1 \text{ is a projection}) \\ &\leq e^{-\mu t}S(Qh) + (1 - e^{-\mu t})S(P_1Qh) \\ &\leq e^{-\mu t}S(h) + (1 - e^{-\mu t})\frac{1}{\binom{N}{2}} \sum_{i < j} S(P_1R_{ij}h) \\ &= e^{-\mu t}S(h) + (1 - e^{-\mu t})\frac{1}{\binom{N}{2}} \left(\sum_{i < j, i, j \neq 1} S(P_1R_{ij}h) + \sum_{j=2}^N S(P_1R_{1j}h) \right) \\ &\leq e^{-\mu t}S(h) + (1 - e^{-\mu t})\frac{1}{\binom{N}{2}} \left(\sum_{i < j, i, j \neq 1} S(h) + (N-1 - \frac{1}{2})S(h) \right) \\ &= \left(1 - \frac{1 - e^{-\mu t}}{N(N-1)}\right) S(h), \end{aligned}$$

where we use Lemma 2.4.10 in the last inequality. We use the convexity of the entropy and the averaging property of P_1 and Q in the previous steps. \square

Lemma 2.4.12. *Let $1 \leq m < N$. Then*

$$S \left(\exp \left(\mu \sum_{k=1}^m (P_k - I)t \right) Qh \right) \leq \left(1 - \frac{m(1 - e^{-\mu t})}{N(N-1)} \right) S(h). \quad (53)$$

Proof. We prove the above by induction on m . The base case $m = 1$ (and any $N > 1$) was shown in the previous Lemma. We restrict to $\{(N, m) : 2 \leq m < N\}$ for the rest of the proof. Assume that the Lemma is true for $m - 1$ (and any $N > m - 1$). To infer from this its validity for the case m (and any $N > m$), we analyze below the entropy of $P_m \exp(\mu \sum_{k=1}^{m-1} (P_k - I)t)$, where we expand the Kac operator Q , split it into terms that contain m and those that do not, and utilize the convexity of the entropy.

$$\begin{aligned} S \left(P_m \exp \left(\mu \sum_{k=1}^{m-1} (P_k - I)t \right) Qh \right) &\leq \left(1 - \frac{2}{N} \right) S \left(\frac{\exp \left(\mu \sum_{k=1}^{m-1} (P_k - I)t \right)}{\binom{N-1}{2}} \sum_{\substack{i < j \\ i, j \neq m}} R_{ij} P_m h \right) \\ &\quad + \frac{2}{N} S \left(\frac{\exp \left(\mu \sum_{k=1}^{m-1} (P_k - I)t \right)}{N-1} P_m \sum_{l \neq m} R_{lm} h \right). \end{aligned}$$

In the first term¹, we also use the commutativity of P_m with R_{ij} when neither i nor j equal m . Next, we treat the terms as follows:

- Term 1: We apply the induction hypothesis for $m - 1$, $N - 1$ since $P_m h$ is a function of $N - 1$ variables and $\binom{N-1}{2}^{-1} \sum_{\substack{i < j \\ i, j \neq m}} R_{ij}$ is the Kac operator acting on $N - 1$ variables.
- Term 2: We use the averaging property of $\exp(\mu \sum_{k=1}^{m-1} (P_k - I)t)$, convexity, and Lemma 2.4.10.

¹This term is non-zero only when $N > 2$, which is the case here.

We obtain

$$S(P_m \exp(\mu \sum_{k=1}^{m-1} (P_k - I)t) Qh) \leq (1 - \frac{2}{N})(1 - (m-1) \frac{1 - e^{-\mu t}}{(N-1)(N-2)}) S(h) \quad (54)$$

$$+ \frac{2}{N} \frac{1}{N-1} (N - \frac{3}{2}) S(h).$$

Now starting with the left-hand side of (53) and using convexity plus the fact that P_m is a projection, write

$$S\left(\exp\left(\mu \sum_{k=1}^m (P_k - I)t\right) Qh\right) = S\left((e^{-\mu t} I + (1 - e^{-\mu t}) P_m) \exp\left(\mu \sum_{k=1}^{m-1} (P_k - I)t\right) Qh\right)$$

$$\leq e^{-\mu t} S\left(\exp\left(\mu \sum_{k=1}^{m-1} (P_k - I)t\right) Qh\right)$$

$$+ (1 - e^{-\mu t}) S\left(P_m \exp\left(\mu \sum_{k=1}^{m-1} (P_k - I)t\right) Qh\right).$$

Using the induction hypothesis for the case $m-1$, N for the first term, and the bound (54) for the second term, the Lemma follows through some algebraic simplification. \square

In the following, denote $A(t) := 1 - \frac{m(1-e^{-\mu t})}{N(N-1)}$.

Proof of Theorem 2.2.5. Expanding $e^{-\mathcal{L}_{N,m}t}$ using the Dyson series with Q as the perturbation:

$$e^{N\lambda(Q-I)t + \mu \sum_k (P_k - I)t} = e^{-N\lambda t} e^{N\lambda Q t + \mu \sum (P_k - I)t}$$

$$= e^{-N\lambda t} \left\{ e^{\mu \sum (P_k - I)t} + \int_0^t dt_1 e^{\mu \sum (P_k - I)(t-t_1)} N\lambda Q e^{\mu \sum (P_k - I)t_1} \right.$$

$$+ \int_0^t dt_1 \int_0^{t_1} dt_2 e^{\mu \sum (P_k - I)(t-t_1)} N\lambda Q e^{\mu \sum (P_k - I)(t_1-t_2)} N\lambda Q e^{\mu \sum (P_k - I)t_2} + \dots \left. \right\}.$$

Therefore, using the convexity of entropy, and Lemma 2.4.12,

$$S(h(., t)) \leq e^{-N\lambda t} \left(1 + N\lambda \int_0^t dt_1 A(t-t_1) + (N\lambda)^2 \int_0^t dt_1 \int_0^{t_1} dt_2 A(t-t_1) A(t_1-t_2) + \dots \right) S(h(., 0))$$

$$= e^{-N\lambda t} (1 + N\lambda(A * 1) + (N\lambda)^2(A * A * 1) + \dots) S(h(., 0)),$$

where $*$ is the Laplace-convolution operation. Thus we have that

$$S(h(., t)) \leq e^{-N\lambda t} \varphi(t) S(h(., 0)) , \quad (55)$$

where φ is defined through the series above. We compute $\varphi(t)$ using its Laplace transform $\tilde{\varphi}(s)$. Then:

$$\tilde{\varphi}(s) = \frac{1}{s} \sum_{k=0}^{\infty} (N\lambda \tilde{A}(s))^k ,$$

where $\tilde{A}(s) = \frac{1}{s} - \frac{m}{N(N-1)}(\frac{1}{s} - \frac{1}{s+\mu})$ is the Laplace transform of $A(t)$.

Summing the geometric series (the sum converges if we assume, for instance, that $\tilde{\varphi}(s)$ is defined on the domain $s > N\lambda$),

$$\tilde{\varphi}(s) = \frac{s + \mu}{s^2 + (\mu - N\lambda)s - N\mu\lambda(1 - \frac{m}{N(N-1)})} .$$

The inverse Laplace transform of the above is

$$-\frac{\xi_- e^{(N\lambda - \xi_+)t}}{\sqrt{(N\lambda + \mu)^2 - 4m\lambda\mu/(N-1)}} + \frac{\xi_+ e^{(N\lambda - \xi_-)t}}{\sqrt{(N\lambda + \mu)^2 - 4m\lambda\mu/(N-1)}} .$$

Now we invoke the uniqueness of the Inverse Laplace Transform: No two piecewise continuous, locally bounded functions of exponential order can have the same Laplace transform (see e.g. [11]). Since $\varphi(t)$ (see eq. (55)) belongs to this space, we get

$$\varphi(t) = -\frac{\xi_- e^{(N\lambda - \xi_+)t}}{\sqrt{(N\lambda + \mu)^2 - 4m\lambda\mu/(N-1)}} + \frac{\xi_+ e^{(N\lambda - \xi_-)t}}{\sqrt{(N\lambda + \mu)^2 - 4m\lambda\mu/(N-1)}} .$$

Plugging this into (55), we obtain the desired result (22). \square

Remarks.

- From (50), one notices that $\frac{dZ}{dt}|_{t=0} = 0$. This implies, in particular, that Theorem 2.2.5 does not give us a bound like (47) on the entropy production for Model I. This results from the fact that the significant bounds used in the proof, from Lemma 2.4.10, required the presence of the second-order term $\sum_k (P_k - I)Q$. Note that $\frac{d^2 Z}{dt^2}|_{t=0} < 0$.

- The main bound (Lemma 2.4.10) was obtained by estimating terms of the form $S(P_1 R_{1j} h)$ with $j > m$ that corresponds to a non-thermostated particle hitting a thermostated one, and we ignored any possible contribution from many other terms e.g. $S(R_{ij} R_{kl} h)$, that arises from “second-order” collisions. Thus, there may be scope for a better bound.
- In particular, we hope to obtain an entropy decay rate that scales as $\frac{m}{N}$ (as we had for the spectral gap). We were able to obtain a decay rate scaling as $\frac{1}{N}$ for a modified model: a system of N particles where one of them is strongly thermostated and the Kac collision interaction is replaced by the (much stronger) projection onto radial functions. Thus, the role of the Kac interaction in the equilibration process needs to be better understood.
- The first remark implies, in particular, that unlike Model I (Subsection 2.4.1), the methods used in this subsection cannot be replicated to obtain a proof in terms of the entropy production. In the case of Model I, unlike here, we were able to show contraction in entropy in the first order term $\frac{1}{N} \sum_{i=1}^N T_i + Q$. At this point, we demonstrate why a common approach to finding an entropy production bound for Model II fails. Consider the case $N = 2$, $m = 1$ with $\lambda = \frac{1}{2}$, $\mu = 1$. Here, one could write

$$\begin{aligned} \frac{dS(h)}{dt} &= \int P_1 h \log h \gamma d\mathbf{v} + \int Q h \log h \gamma d\mathbf{v} - 2S(h) \\ &\leq \int P_1 h \log P_1 h \gamma d\mathbf{v} + \int Q h \log Q h \gamma d\mathbf{v} - 2S(h) . \end{aligned}$$

We use in the last step that P_1, Q are projections and $\log x$ is concave. Bounding this from above by $-kS(h)$ (for some $k > 0$) would be sufficient to obtain an entropy production bound. This idea has worked, e.g., for a sum of mutually orthogonal projections like strong thermostats P acting on different particles.

However, in our case, we can find, for every $\epsilon > 0$, a density h_ϵ such that

$$\frac{\int P_1 h_\epsilon \log P_1 h_\epsilon \gamma d\mathbf{v} + \int Q h_\epsilon \log Q h_\epsilon \gamma d\mathbf{v}}{S(h_\epsilon)} \geq 2 - \epsilon . \quad (56)$$

The details are shown in Appendix B. The idea is to take h proportional to the characteristic function of the set $[-a, a] \times [R - a, R + a]$. As $R \rightarrow \infty$, the ratio above asymptotically approaches the value 2. The intuition behind this construction is that as $R \rightarrow \infty$, h is supported approximately in the intersection of the supports of $P_1 h$ (a “band” of width $2a$ parallel to the v_1 axis) and Qh (an annulus around the origin). It is the tangential nature of this intersection that precludes the application of Han’s inequality [18] to improve the bound $S(P_1 h) + S(Qh) \leq 2S(h)$. We are not, however, ruling out the possibility of using a different method to obtain an entropy production bound.

2.5 Approach to Equilibrium in the Gabetta-Toscani-Wennberg Distance

Given two probability densities $f_1(\mathbf{v}), f_2(\mathbf{v})$, the distance $d_2(f_1, f_2)$ introduced by Gabetta, Toscani, Wennberg [14] is defined as

$$d_2(f_1, f_2) := \sup_{\vec{\xi} \neq \mathbf{0}} \frac{|\hat{f}_1(\vec{\xi}) - \hat{f}_2(\vec{\xi})|}{|\vec{\xi}|^2} , \quad (57)$$

where \hat{f}_1, \hat{f}_2 denote the Fourier transforms of f_1, f_2 respectively. That is,

$$\hat{f}(\vec{\xi}) := \int f(\mathbf{v}) e^{-2\pi i \vec{\xi} \cdot \mathbf{v}} d\mathbf{v} .$$

We will call the above metric the GTW distance.

Theorem 2.5.1. *In Theorems 2.2.4 and 2.2.5, $S(f|\gamma)$ can be replaced by $d_2(f, \gamma)$.*

The Theorem follows from the following lemmas that essentially state that $d_2(f, \gamma)$ satisfies those properties of the relative entropy $S(f|\gamma)$ that are needed in the proofs of the Theorems showing equilibration in entropy.

The first is the convexity of $d_2(f, \gamma)$ in each argument, similar to Lemma 2.4.1 for the entropy. In particular,

Lemma 2.5.2. *Let $a, b \geq 0$ such that $a + b = 1$. Let f_1, f_2 be probability densities on \mathbb{R}^N . Then*

$$d_2(af_1 + bf_2, \gamma) \leq ad_2(f_1, \gamma) + bd_2(f_2, \gamma) .$$

Proof. This follows directly from the definition of d_2 , writing $\hat{\gamma}$ as $a\hat{\gamma} + b\hat{\gamma}$. \square

The next Lemma shows the contractive property of the GTW distance under the thermostat, similar to Proposition 2.4.3:

Lemma 2.5.3. *Let $f(v)$ be a probability density on \mathbb{R} , and $g(v)$ the Gaussian. Then*

$$d_2(Wf, g) \leq \frac{1}{2}d_2(f, g) .$$

Proof. Starting with the left-hand side,

$$\begin{aligned} d_2(Wf, g) &= \sup_{\xi \neq 0} \frac{|\widehat{Wf}(\xi) - \hat{g}(\xi)|}{\xi^2} \\ &= \sup_{\xi \neq 0} \frac{|\widehat{Wf}(\xi) - \widehat{Wg}(\xi)|}{\xi^2} \\ &= \sup_{\xi \neq 0} \frac{|f \, d\theta \left(\hat{f}(\xi \cos \theta) - \hat{g}(\xi \cos \theta) \right) \hat{g}(\xi \sin \theta)|}{\xi^2} \\ &= \sup_{\xi \neq 0} \frac{|f \, d\theta \left(\hat{f}(\xi \cos \theta) - \hat{g}(\xi \cos \theta) \right) \hat{g}(\xi \sin \theta) \cos^2 \theta|}{\xi^2 \cos^2 \theta} \\ &\leq d_2(f, g) \left(\int d\theta \cos^2 \theta \right) \quad (\text{since } |\hat{g}| \leq 1) \\ &= \frac{1}{2}d_2(f, g) . \end{aligned}$$

\square

Now we show a version of Han's Inequality for the GTW distance, similar to Lemma 2.4.6 for the relative entropy.

Lemma 2.5.4. *Let $f(\mathbf{v})$ be a probability density on \mathbb{R}^N and let $f_i := \int dv_i f(v_1, \dots, v_i, \dots, v_N)$ be the i -th particle marginal. Then*

$$\sum_{i=1}^N d_2(f_i, \gamma) \leq (N-1)d_2(f, \gamma) .$$

Proof. Note that the Fourier transform of f_i is equal to $\hat{f}|_{\xi_i=0}$. The left-hand side is

$$\begin{aligned} \sum_{i=1}^N d_2(f_i, \gamma) &= \sum_{i=1}^N \sup_{\vec{\xi}} \frac{|\hat{f}(\xi_1, \dots, \xi_i = 0, \dots, \xi_N) - \hat{\gamma}|}{|\vec{\xi}|^2} \\ &= \sum_{i=1}^N \sup_{\vec{\xi}} \frac{|\hat{f}(\xi_1, \dots, \xi_i = 0, \dots, \xi_N) - \hat{\gamma}|}{|\vec{\xi}|^2} \times \left(\frac{|\xi_1^2 + \dots + 0 + \dots + \xi_N^2|}{|\xi_1^2 + \dots + 0 + \dots + \xi_N^2|} \right) \\ &\leq d_2(f, \gamma) \sum_{i=1}^N \frac{|\vec{\xi}|^2 - \xi_i^2}{|\vec{\xi}|^2} \\ &= (N-1)d_2(f, \gamma) . \end{aligned}$$

□

Like the relative entropy (Lemma 2.4.8), the distance $d_2(f, \gamma)$ also exhibits a weaker version of the previous Lemma in the case of a sum of thermostats W_i , instead of the strong thermostats P_i .

Lemma 2.5.5.

$$d_2\left(\frac{1}{N} \sum_{i=1}^N W_i f, \gamma\right) \leq \left(\frac{N - \frac{1}{2}}{N}\right) d_2(f, \gamma) .$$

Proof.

$$\begin{aligned} d_2\left(\frac{1}{N} \sum_{i=1}^N W_i f, \gamma\right) &\leq \sup_{\vec{\xi}} \frac{1}{|\vec{\xi}|^2} \frac{1}{N} \sum_{i=1}^N \left| \int d\theta \left(\hat{f}(\xi_1, \dots, \xi_i \cos \theta, \dots, \xi_N) - \hat{g}(\xi_i \cos \theta) \prod_{j \neq i} \hat{g}(\xi_j) \right) \hat{g}(\xi_i \sin \theta) \right| \\ &\leq d_2(f, \gamma) \sup_{\vec{\xi}} \frac{1}{|\vec{\xi}|^2} \frac{1}{N} \sum_{i=1}^N \int d\theta (\xi_1^2 + \dots + \xi_i^2 \cos^2 \theta + \dots + \xi_N^2) \\ &= d_2(f, \gamma) \frac{1}{N} \left(N - \frac{1}{2} \right) . \end{aligned}$$

□

The final property we need is the adaptation of Lemma 2.4.2, the non-expansivity of the GTW distance from equilibrium under the Kac operator Q .

Lemma 2.5.6.

$$\sup_{\vec{\xi} \neq 0} \frac{|\widehat{Qf}(\vec{\xi})|}{|\vec{\xi}|^2} \leq \sup_{\vec{\xi} \neq 0} \frac{|\hat{f}(\vec{\xi})|}{|\vec{\xi}|^2} .$$

Proof.

$$\begin{aligned} \sup_{\vec{\xi} \neq 0} \frac{|\widehat{Qf}(\vec{\xi})|}{|\vec{\xi}|^2} &\leq \frac{1}{\binom{N}{2}} \sum_{i < j} \sup_{\vec{\xi} \neq 0} \frac{|\hat{f}(\dots, \xi_i^*, \dots, \xi_j^*, \dots)|}{|\vec{\xi}|^2} \\ &= \frac{1}{\binom{N}{2}} \sum_{i < j} \sup_{\vec{\xi} \neq 0} \frac{|\hat{f}(\dots, \xi_i^*, \dots, \xi_j^*, \dots)|}{|\xi_1^2 + \dots + \xi_i^{*2} + \dots + \xi_j^{*2} + \dots + \xi_N^2|} = \sup_{\vec{\xi} \neq 0} \frac{|\hat{f}(\vec{\xi})|}{|\vec{\xi}|^2} . \end{aligned}$$

□

2.6 Macroscopic Limit

In order to study the $N \rightarrow \infty$ limit of Model I, it is interesting to see how the notion of propagation of chaos (Section 1.2) looks in our scenario. Since the probability distribution is assumed to be symmetric in each v_i , the primary difference from Definition 1.2.1 of a chaotic sequence is that the phase space is now \mathbb{R}^N .

Given a distribution $f^{(N)}(\mathbf{v})$ with $\mathbf{v} \in \mathbb{R}^N$, we can define the k particle marginal as

$$f_k^{(N)}(v_1, \dots, v_k) = \int f^{(N)}(\mathbf{v}) \prod_{i=k+1}^N dv_i .$$

Definition 2.6.1. A sequence of probability distributions $\{f^{(N)}(\mathbf{v})\}_{N=1}^\infty$ on \mathbb{R}^N is said to be **chaotic** if, $\forall k \geq 1$, we have

$$\lim_{N \rightarrow \infty} f_k^{(N)}(v_1, \dots, v_k) = \lim_{N \rightarrow \infty} \prod_{j=1}^k f_1^{(N)}(v_j) ,$$

where the above limit is taken in the weak sense.

Given that at time $t = 0$, we have a chaotic sequence of distributions, the evolution equation (11) can be shown to preserve this property for all t , through a simple generalization of the proof by Kac [21] (see also [26]). The statement, and proof idea of the *propagation of chaos* statement are below.

Theorem 2.6.2. *Let $f^{(N)}(\mathbf{v}, 0)$ be a chaotic sequence of initial densities. Then its evolution under the master equation (11), $f^{(N)}(\mathbf{v}, t)$, is a chaotic sequence for any fixed t . That is, if*

$$\lim_{N \rightarrow \infty} \int_{\mathbb{R}^N} \varphi_1(v_1) \dots \varphi_k(v_k) f^{(N)}(\mathbf{v}, 0) = \prod_{j=1}^k \lim_{N \rightarrow \infty} \int_{\mathbb{R}^N} \varphi_j(v_j) f^{(N)}(\mathbf{v}, 0)$$

for any $k \in \mathbb{N}$ and any $\phi_1(v_1), \dots, \phi_k(v_k)$ bounded and continuous, then for any t :

$$\lim_{N \rightarrow \infty} \int_{\mathbb{R}^N} \varphi_1(v_1) \dots \varphi_k(v_k) f^{(N)}(\mathbf{v}, t) = \prod_{j=1}^k \lim_{N \rightarrow \infty} \int_{\mathbb{R}^N} \varphi_j(v_j) f^{(N)}(\mathbf{v}, t)$$

for any $k \in \mathbb{N}$ and any $\varphi_1(v_1), \dots, \varphi_k(v_k)$ bounded and continuous.

We adapt McKean's version [26] of Kac's proof [21]. The idea is to write $f(\mathbf{v}, t) = e^{-(\lambda \mathcal{G}_K + \mu \mathcal{G}_T)t} f(\mathbf{v}, 0)$, expand the exponential in series of t , and use the chaotic property of the initial sequence. The key observation is that \mathcal{G}_T is a derivation already for finite N (in the sense of Lemma 2.6.4). Two main ingredients are needed:

Lemma 2.6.3. *The series $\sum_{l=0}^{\infty} \frac{t^l}{l!} \int \varphi_1(v_1) \dots \varphi_k(v_k) (\lambda \mathcal{G}_K + \mu \mathcal{G}_T)^l f(\mathbf{v}, 0)$ converges absolutely if $t < \frac{1}{4\lambda + \mu}$.*

Proof. To prove the lemma, it is enough to show that:

$$\|(\lambda \mathcal{G}_K + \mu \mathcal{G}_T^*)^l \phi\|_{\infty} \leq (4\lambda + 2\mu)^l m(m+1) \dots (m+l-1) \|\phi\|_{\infty} \quad (58)$$

and then follow the proof in [26]. The above statement follows from a simple induction starting from

$$|(\lambda \mathcal{G}_K + \mu \mathcal{G}_T^*) \phi(v_1, \dots, v_m)| \leq |\lambda \mathcal{G}_K \phi| + |\mu \mathcal{G}_T^* \phi| \leq (4\lambda + 2\mu)m \|\phi\|_{\infty}.$$

□

Calling

$$\Gamma_K \phi := 2 \sum_{i \leq m} \oint d\theta (\phi(\dots, v_i \cos \theta + v_{m+1} \sin \theta, \dots) - \phi),$$

one can prove, as in [26], that if $\varphi_1(v_1), \dots, \varphi_k(v_k)$ are bounded and continuous then:

$$\lim_{N \rightarrow \infty} \int (\lambda \mathcal{G}_K + \mu \mathcal{G}_T^*)^l [\varphi_1 \dots \varphi_k] f^{(N)}(\mathbf{v}, 0) = \lim_{N \rightarrow \infty} \int (\lambda \Gamma_K + \mu \mathcal{G}_T^*)^l [\varphi_1 \dots \varphi_k] f^{(N)}(\mathbf{v}, 0) .$$

The main ingredient to re-sum the power series expansion and obtain the Boltzmann equation is the following “algebraic” Lemma.

Lemma 2.6.4. *If $(\phi \otimes \psi)(v_1, \dots, v_{m+k}) := \phi(v_1, \dots, v_m) \psi(v_{m+1}, \dots, v_{m+k})$, then*

$$(\Gamma_K + \mathcal{G}_T^*)[\phi \otimes \psi] = (\Gamma_K + \mathcal{G}_T^*)[\phi] \otimes \psi + \phi \otimes (\Gamma_K + \mathcal{G}_T^*)[\psi] .$$

It is now possible to prove Theorem 2.2.6 by following the proof in [26] step-by-step. Thus, in the macroscopic limit $N \rightarrow \infty$, all the information is contained in the single particle marginal that satisfies the Boltzmann equation in 2.2.6.

We close the discussion on Model I with a few comments on the “Boltzmann equation” in Theorem 2.2.6. The equilibrium is the Maxwellian $\frac{1}{2\pi} e^{-v^2/2}$, and the relative entropy along the evolution with respect to this decays at an exponential rate $\frac{\mu}{2}$, which can be shown to be optimal (these follow from the results of Subsection 2.4.1). The evolution operator resulting from the *linearized* version of the Boltzmann equation is diagonalized by the Hermite polynomials (both the collision and thermostat parts), with the n -th degree polynomial $H_n(v)$ yielding eigenvalues $2\lambda(1 - 2s_n)$ and $\mu(1 - s_n)$, respectively, where $s_n := \int_0^{2\pi} \cos^n \theta d\theta$. Thus, the gap is $\frac{\mu}{2}$, and the “second” gap is $\frac{\lambda}{2} + \frac{5}{8}\mu$, which correspond to the $N \rightarrow \infty$ limit of the respective gaps found at the Master equation level (Proposition 2.2.1 and Theorem 2.2.2). Incidentally, the eigenvalue μ found in the latter (see 18) does not appear here since the single-particle marginal of the corresponding eigenfunction vanishes in the limit.

In Model II, when m is finite, the coupling to the heat bath becomes insignificant in the thermodynamic limit. On the other hand, when $m = \alpha N$ for some $\alpha < 1$, we expect that a coupled Boltzmann equation system should result. The *Stosszahlansatz* needs to be reformulated in a precise manner to adapt to our situation where there is

an asymmetry between the strongly thermostated and the non-thermostated particles. Moreover, generalizations of our model could bring about connections to previously studied thermostated Boltzmann equations [9].

2.7 Measures of Equilibration

In the previous section, we have used three indicators of equilibration: the spectral gap, relative entropy and the GTW distance. Here, we compare their strengths and weaknesses as measures of equilibration. The kinetic energy, although a physically interesting quantity, is clearly not a very strong indicator, and we leave it out of the current discussion. Apart from general comments, we compare how each of these indicators act on distributions that are approximately independent: $F(\mathbf{v}) \sim f(v_1) \dots f(v_N)$ that arise in many physical scenarios.

The spectral gap Δ indicates an exponential decay in the L^2 distance as follows:

$$\left\| \frac{F(\mathbf{v}, t)}{\gamma} - 1 \right\|_{L^2(\mathbb{R}^N, \gamma)} \leq e^{-\Delta t} \left\| \frac{F(\mathbf{v}, 0)}{\gamma} - 1 \right\|_{L^2(\mathbb{R}^N, \gamma)} .$$

If the initial distribution is a product, $\left\| \frac{F(\mathbf{v}, 0)}{\gamma} - 1 \right\|_{L^2(\mathbb{R}^N, \gamma)} \sim \prod_j \|f(v_j)\|_{L^2(\mathbb{R})} \sim c^N$, for $c > 1$. Hence, one needs to wait for a time order N to observe equilibration even if Δ remains strictly positive independent of N . This is not favourable. Moreover, as mentioned in the previous section, the spectral gap information can be gleaned from the linearized Boltzmann equation. Nevertheless, studying the L^2 distance is a good idea in general as i) the analysis lends itself to tools from spectral theory ii) obtaining the eigenfunctions in explicit form yields more insight into the equilibration process iii) it indicates behavior close to equilibrium. In the discussion following eq. (47) for Model I, we saw that the spectral gap places an upper bound on the entropy production bound. In Model II, where we did not have an entropy production bound, the gap helped understand the “lowest-order” behavior.

The relative entropy has historical significance and a physical meaning in equilibrium. And importantly, it is an extensive property, and hence a good measure of equilibration, independent of the dimension of the phase space.

$$S(\prod_j f(v_j)|\gamma) = \sum_j S(f(v_j)|\gamma(v_j))$$

The GTW distance also behaves well on product distributions. The triangle inequality can be employed to show that d_2 is sub-extensive.

$$d_2(\prod_j f(v_j), \gamma) \leq \sum_j d_2(f(v_j), \gamma(v_j))$$

In addition, it can be used to prove existence, uniqueness and approach to the steady state in circumstances when the steady state itself is not explicitly known [9].

CHAPTER III

THE THERMOSTAT MODEL: TWO PERSPECTIVES

We look deeper into the thermostat model used in Model I. The rationale behind the model, as seen before, is as follows: given a test function $\phi(\mathbf{v})$ of the state space, and an initial probability distribution $f(\mathbf{v})$, the expectation value of ϕ after Particle j collides with the thermostat is

$$\int d\mathbf{v} dw g(w) \oint d\theta \phi(\dots, v_{j-1}, v_j \cos \theta + w \sin \theta, v_{j+1}, \dots) f(\mathbf{v}) ,$$

which corresponds to a modification in the system state space due to a Kac style collision with a particle with the Gaussian velocity distribution $g(w)$. The thermostat is assumed to be *ideal*: its state itself is not altered by interaction with the system.

In this chapter, we discuss two approaches to simulating the action of this thermostat on a particle dynamically, as a limit of systems with non-ideal thermostating: i) By letting the particle undergo Kac collisions with a huge number of particles with the Gaussian velocity distribution (Section 3.1) and ii) By letting the particle undergo Kac collisions with a single particle whose velocity distribution is “reset” to the Gaussian at a huge rate (Section 3.2). We describe the two treatments and prove quantitative statements in the following two sections. The main results are Theorems 3.1.12, 3.1.13 and Theorem 3.2.1. The results in Section 3.2 have been published in [30].

3.1 A Microscopic Realization

In this section, we show in terms of the GTW and L^2 metrics that the thermostat model can be realized as a tagged particle colliding via the Kac mechanism with

an infinite heat bath. Consider Particle 1 (the tagged particle) with initial velocity distribution $f(v_1)$ interacting with $(N - 1)$ particles $2, \dots, N$ whose velocities have the equilibrium Gaussian distribution. Then we show that for every $t \geq 0$,

$$e^{N(Q-I)t} (f(v_1)g(v_2)\dots g(v_N)) \longrightarrow (e^{2(W_1-I)t} f(v_1)) g(v_2)\dots g(v_N)$$

as $N \rightarrow \infty$ *uniformly* in t . Notice the placement of the parentheses above: i) the right hand side evolution does not change the state of particles $2, \dots, N$, and ii) the Kac interaction on the left hand side mixes the states of all particles. Particles $2, \dots, N$ are perturbed from the Gaussian state due to their mixing with Particle 1. However, this perturbation is spread across $N - 1$ particles, and turns out to be negligible as $N \rightarrow \infty$.

To start with, we present a preliminary computation on the above limit in the special case of $t \rightarrow \infty$. In this limit, the Kac evolution leads the system probability distribution to the radial projection of the initial condition (recall eq. (25))

$$\oint_{S^{N-1}(|\mathbf{v}|)} d\sigma(\mathbf{w}) f(w_1)g(w_2)\dots g(w_N) =: Bf$$

whereas the thermostat evolution (which only acts on Particle 1) takes it to the full Gaussian $\gamma(\mathbf{v})$, and thus with no memory of the initial state.

Starting with the observation that the Gaussian is a radial function, and that the Fourier transform \mathcal{F} commutes with B , the GTW distance between Bf and γ is given by

$$\begin{aligned} d_2(Bf, \gamma) &= \sup_{\vec{\xi} \neq 0} \frac{|\mathcal{F}[B(f - \gamma)](\vec{\xi})|}{|\vec{\xi}|^2} \\ &= \sup_{\vec{\xi} \neq 0} \frac{|B[(\hat{f}(\xi_1) - \hat{g}(\xi_1)) \prod_{j=2}^N \hat{g}(\xi_j)]|}{|\vec{\xi}|^2} \\ &\leq \sup_{\rho \neq 0} \frac{1}{\rho^2} \oint_{\eta \in S^{N-1}(\rho)} |\hat{f}(\eta_1) - \hat{g}(\eta_1)| \prod_{j=2}^N \hat{g}(\eta_j) \\ &\leq d_2(f, g) \sup_{\rho \neq 0} \frac{1}{\rho^2} \oint \eta_1^2 \end{aligned}$$

$$= \frac{1}{N} d_2(f, g) .$$

On the other hand, the L^2 distance is measured (as in Chapter 2) by considering the norm of the function $\frac{f(v_1)g(v_2)\dots g(v_N)}{g(v_1)\dots g(v_N)} - 1$ in space $L^2(\mathbb{R}^N, \gamma(\mathbf{v})d\mathbf{v})$. Define $u(v_1) := \frac{f(v_1)}{g(v_1)}$ and note that $\langle u - 1, 1 \rangle = 0$, where the inner product is in the aforementioned space. We write, without loss of generality,

$$u - 1 = \sum_{k=1}^{\infty} c_k H_{2k}(v_1) ,$$

where H_{2k} are Hermite polynomials of degree $2k$ as defined in Chapter 2. Notice that H_0 is not in the sum, since $u - 1$ is orthogonal to 1. We can restrict to even Hermite polynomials as the odd ones lie in the kernel of B so that $B(u - 1) = 0$, and there is nothing to show.

$$B(u - 1) = \sum_{k=1}^{\infty} c_k B(H_{2k})$$

Now, from eq. (28) and the fact that B satisfies the hypotheses of Lemma 2.3.7,

$$BH_{2k}(v_1) = \Gamma(k, 0, \dots, 0) \sum_{|\underline{\beta}|=k} \frac{k!}{\beta_1! \dots \beta_N!} H_{2\beta_1}(v_1) \dots H_{2\beta_N}(v_N) . \quad (59)$$

Observing that BH_{2k} is a linear combination of Hermite polynomials of total degree $2k$, we get from the orthogonality of the Hermite polynomials that

$$\|B(u - 1)\|^2 = \sum_{k=1}^{\infty} |c_k|^2 \|BH_{2k}\|^2 . \quad (60)$$

Now, we compute $\|BH_{2k}\|^2 = \langle BH_{2k}, BH_{2k} \rangle = \langle BH_{2k}, H_{2k} \rangle$ (B is a self-adjoint projection in $L^2(\mathbb{R}^N, \gamma(\mathbf{v})d\mathbf{v})$) to get from eq. (59) that

$$\|BH_{2k}(v_1)\|^2 = \Gamma(k, 0, \dots, 0) \|H_{2k}(v_1)\|^2 .$$

Finally, it is easy to see that $\Gamma(k, 0, \dots, 0) = \frac{\Gamma(\frac{N}{2})\Gamma(k+\frac{1}{2})}{\Gamma(\frac{N}{2}+k)} \leq \frac{\sqrt{\pi}}{N}$ when $N \geq 2$, and thus

$$\|B(u - 1)\|^2 \leq \frac{\sqrt{\pi}}{N} \sum_{k=1}^{\infty} |c_k|^2 \|H_{2k}(v_1)\|^2 = \frac{\sqrt{\pi}}{N} \|u - 1\|^2 .$$

In other words, $\|B(u-1)\| \leq \frac{\text{const.}}{\sqrt{N}} \|u-1\|$.

The above computations give us an idea of what quantitative bounds (in terms of N) to expect for the general t case. Indeed, we will state and prove our result for general t below, where the scaling is seen to be $\frac{1}{N}$ for the d_2 metric, and $\frac{1}{\sqrt{N}}$ for the L^2 metric.

An analysis of the limit to be studied yields

$$\begin{aligned}
& (e^{Nt(Q-I)} - e^{2(W_1-I)t}) f(v_1) \prod_{j=2}^N g(v_j) = (e^{Nt(Q-I)} - e^{2(W_1-I)t}) f(v_1) \prod_{j=2}^N g(v_j) \\
& = e^{-Nt} \sum_{k=0}^{\infty} \frac{N^k t^k}{k!} \left[Q^k - \left(1 - \frac{2}{N} + \frac{2}{N} W_1\right)^k \right] f(v_1) \prod_{j=2}^N g(v_j) \\
& = e^{-Nt} \sum_{k=0}^{\infty} \frac{N^k t^k}{k!} \sum_{m=0}^{k-1} Q^{k-1-m} \left(Q - \left(1 - \frac{2}{N} + \frac{2}{N} W_1\right) \right) \left(1 - \frac{2}{N} + \frac{2}{N} W_1\right)^m f(v_1) \prod_{j=2}^N g(v_j) .
\end{aligned} \tag{61}$$

Since $\left(1 - \frac{2}{N} + \frac{2}{N} W_1\right)^m f(v_1) \prod_{j=2}^N g(v_j)$ is a function of v_1 multiplied by a Gaussian in variables v_2, \dots, v_N in every term in the sum, the action of the operator

$$Q - \left(1 - \frac{2}{N} + \frac{2}{N} W_1\right)$$

on such a function simplifies to

$$\frac{1}{\binom{N}{2}} \left(\sum_{j=2}^N R_{1j} + \binom{N-1}{2} \right) - \left(1 - \frac{2}{N} + \frac{2}{N} W_1\right) = \frac{2}{N} \left(\frac{1}{N-1} \sum_{j=2}^N R_{1j} - W_1 \right) . \tag{62}$$

We will use two different representations of the initial state $f(v_1)$:

- For the L^2 distance, consider $u(v) = \frac{f(v)}{g(v)} \in L^2(\mathbb{R}, g(v)dv)$, with $\int u(v)g(v)dv = 0$. The operator Q does not change since $Q[\frac{f(v)}{g(v)}] = \frac{Qf(v)}{g(v)}$ but W_1 transforms into T_1 , defined as

$$T_1 u(v) := \int dw g(w) \oint d\theta u(v \cos \theta + w \sin \theta) .$$

It can easily be seen that $W_1[f(v)g(v)] = [T_1 f(v)]g(v)$.

- $h(v) = f(v) - g(v)$, for the result in terms of the GTW distance. Neither operators Q or W_1 change in this representation. We have naturally that $h \in L^1(\mathbb{R})$ and $\int h(v)dv = 0$.

Inspired by eq. (62), the following two main Lemmas investigate the relationship between the operators $\frac{1}{N-1} \sum_{j=2}^N R_{1j}$ and the thermostat on the first particle, acting on functions that are either independent of v_2, \dots, v_N (for the L^2 case) or rotationally invariant in v_2, \dots, v_N (for the GTW metric case). As $N \rightarrow \infty$, the two operators are close, quantifying the idea that the thermostat is formed out of repeated Kac collisions of Particle 1 with Gaussians.

Lemma 3.1.1. *Let $u \in L^2(\mathbb{R}, g(v)dv)$. Then*

$$\left\| \left(\frac{1}{N-1} \sum_{j=2}^N R_{1j} - T_1 \right) u(v_1) \right\|^2 = \frac{1}{N-1} \langle T_1 u, (I - T_1) u \rangle ,$$

where the inner product and norm above are in $L^2(\mathbb{R}^N, \gamma(\mathbf{v})d\mathbf{v})$.

Proof. First, we recall the definition $P_j f(v_1, \dots, v_N) := \int dv_j g(v_j) f(\dots, v_j, \dots)$, and note the relationship between the thermostat and the strong thermostat:

$$P_j R_{ij} P_j = P_j T_i . \quad (63)$$

$$\begin{aligned} \left\| \left(\frac{1}{N-1} \sum_{j=2}^N R_{1j} - T_1 \right) u(v_1) \right\|^2 &= \frac{1}{(N-1)^2} \sum_{j=2}^N \langle R_{1j} u, u \rangle + \frac{1}{(N-1)^2} \sum_{j \neq k} \langle R_{1j} u, R_{1k} u \rangle \\ &\quad - \frac{2}{N-1} \sum_{j=2}^N \langle R_{1j} u, T_1 u \rangle + \langle T_1 u, T_1 u \rangle , \end{aligned}$$

where we use that each R_{1j} is a projection on $L^2(\mathbb{R}, \gamma d\mathbf{v})$, and hence self-adjoint and idempotent. Next, we use repeatedly the fact that u is independent of v_j for $j \geq 2$ to add P_j and P_k at various points without loss of generality:

$$\left\| \left(\frac{1}{N-1} \sum_{j=2}^N R_{1j} - T_1 \right) u(v_1) \right\|^2 = \frac{1}{(N-1)^2} \sum_{j=2}^N \langle P_j R_{1j} P_j u, u \rangle$$

$$\begin{aligned}
& + \frac{1}{(N-1)^2} \sum_{j \neq k} \langle P_k R_{1j} P_j u, P_j R_{1k} P_k u \rangle - \frac{2}{N-1} \sum_{j=2}^N \langle R_{1j} P_j u, P_j T_1 u \rangle + \langle T_1 u, T_1 u \rangle \\
& = \frac{1}{(N-1)^2} \sum_{j=2}^N \langle P_j T_1 u, u \rangle + \frac{1}{(N-1)^2} \sum_{j \neq k} \langle P_j T_1 u, P_k T_1 u \rangle \\
& \quad - \frac{2}{N-1} \sum_{j=2}^N \langle P_j T_1 u, T_1 u \rangle + \langle T_1 u, T_1 u \rangle ,
\end{aligned}$$

where the last step follows from (63). Finally, getting rid of the P_j and P_k since $T_1 u$ is a function of v_1 alone, we get the desired result. \square

Next, we state and prove an analogous result in the GTW distance d_2 , for which we introduce the following functionals.

Definition 3.1.2. Let $h \in L^1(\mathbb{R})$ with $\int h(v) dv = 0$, $h(-v) = h(v)$, $\int v |h(v)| dv < \infty$ and $\int v^2 |h(v)| dv < \infty$. Define

$$\mathcal{D}(h) := \sup_{\xi, \eta \neq 0} \frac{|\hat{h}(\sqrt{\xi^2 + \eta^2}) - \hat{h}(|\xi|)|}{\eta^2} \quad (64)$$

$$d_1(h) = \sup_{\xi \neq 0} \frac{|\hat{h}(\xi)|}{|\xi|} \quad (65)$$

where the Fourier Transform $\hat{h} = \mathcal{F}(h) = \int e^{-2\pi i \xi \cdot v} h(v) dv$.

The following Lemma shows that the assumptions on h guarantee the finiteness of $\mathcal{D}(h)$.

Lemma 3.1.3. Let $\varphi \geq 0$ in L^1 be such that $\int \varphi(v) dv = 1$ and $\varphi(-v) = \varphi(v)$. Let ν_φ be its variance and $\hat{\varphi}$ be its Fourier Transform. Then

$$|\hat{\varphi}(\xi_1) - \hat{\varphi}(\xi_2)| \leq 2\pi^2 \nu_\varphi |\xi_1^2 - \xi_2^2| .$$

Proof.

$$\begin{aligned}
|\hat{\varphi}(\xi_1) - \hat{\varphi}(\xi_2)| & \leq \int dv |\cos(2\pi \xi_1 v) - \cos(2\pi \xi_2 v)| \varphi(v) \\
& \leq \int dv \varphi(v) |2 \sin[\pi(\xi_1 + \xi_2)v] \sin[\pi(\xi_2 - \xi_1)v]| \leq 2\pi^2 \int dv \varphi(v) |\xi_1^2 - \xi_2^2| v^2 ,
\end{aligned}$$

which proves the Lemma. \square

We will use the above Lemma in particular, with $\varphi(v) = g(v)$, the Gaussian, to obtain bounds. We now have the main Lemma in the case of the GTW distance:

Lemma 3.1.4. *Let h satisfy the assumptions in Definition 3.1.2. Then*

$$d_2 \left(\frac{1}{N-1} \sum_{i=2}^N R_{1i} h(v_1) \prod_{j=2}^N g(v_j), W_1 h(v_1) \prod_{j=2}^N g(v_j) \right) \leq \frac{1}{N-1} \left(\frac{1}{2} \mathcal{D}(h) + d_1(h) \right).$$

Proof.

$$\begin{aligned} & d_2 \left(\frac{1}{N-1} \sum_{i=2}^N R_{1i} h(v_1) \prod_{j=2}^N g(v_j), W_1 h(v_1) \prod_{j=2}^N g(v_j) \right) \leq \\ & \sup_{\vec{\xi}} \frac{1}{|\vec{\xi}|^2} \frac{1}{N-1} \left| \sum_{i=2}^N \left[\int d\theta \hat{h}(\sqrt{\xi_1^2 + \xi_i^2} \cos \theta) \hat{g}(\sqrt{\xi_1^2 + \xi_i^2} \sin \theta) - \int d\theta \hat{h}(\xi_1 \cos \theta) \hat{g}(\xi_1 \sin \theta) \hat{g}(\xi_i) \right] \right| \\ & \leq \sup_{\vec{\xi}} \frac{1}{|\vec{\xi}|^2} \frac{1}{N-1} \sum_{i=2}^N \left| \int d\theta [\hat{h}(\sqrt{\xi_1^2 + \xi_i^2} \cos \theta) - \hat{h}(\xi_1 \cos \theta)] \hat{g}(\sqrt{\xi_1^2 + \xi_i^2} \sin \theta) \right| \\ & \quad + \left| \int d\theta \hat{h}(\xi_1 \cos \theta) [\hat{g}(\xi_1 \sin \theta) \hat{g}(\xi_i) - \hat{g}(\sqrt{\xi_1^2 + \xi_i^2} \sin \theta)] \right| \\ & \leq \sup_{\vec{\xi}} \frac{1}{N-1} \sum_{i=2}^N \frac{\xi_i^2}{|\vec{\xi}|^2} \frac{1}{\xi_i^2} \left[\int d\theta \mathcal{D}(h) \xi_i^2 \cos^2 \theta + 2\pi^2 \left(\frac{1}{2\pi} \right) \int d\theta \left| \hat{h}(\xi_1 \cos \theta) \right| \hat{g}(\xi_1 \sin \theta) \xi_i^2 \cos^2 \theta \right] \end{aligned}$$

(where the second term in the previous step is obtained by applying Lemma 3.1.3 to the Gaussian. The variance $\nu_g = \frac{1}{2\pi}$.)

$$\begin{aligned} & \leq \frac{1}{N-1} \left[\frac{1}{2} \mathcal{D}(h) + d_1(h) \pi \sup_{\xi} \int d\theta |\xi_1| |\cos^3 \theta| e^{-\pi \xi_1^2 \sin^2 \theta} \right] \\ & \leq \frac{1}{N-1} \left[\frac{1}{2} \mathcal{D}(h) + d_1(h) \right]. \end{aligned}$$

The last step is obtained as follows. Set $w = |\xi_1| \sin \theta$ to get

$$\int d\theta |\xi_1| |\cos^3 \theta| e^{-\pi \xi_1^2 \sin^2 \theta} = \frac{2}{\pi} \int_0^{|\xi_1|} dw \left(1 - \left(\frac{w}{|\xi_1|} \right)^2 \right) e^{-\pi w^2} \leq \frac{2}{\pi} \int_0^\infty e^{-\pi w^2} dw = \frac{1}{\pi}.$$

□

The next few Lemmas investigate the behavior of the term $(1 - \frac{2}{N} + \frac{2}{N} W_1)^m f(v_1) \prod_{j=2}^N g(v_j)$ that appears in eq. (61).

Lemma 3.1.5. *Let $u \in L^2(\mathbb{R}, g(v)dv)$ be such that $\langle u, 1 \rangle = 0$. Then*

$$\langle T_1^2 u, (I - T_1)u \rangle \leq \frac{1}{2} \langle T_1 u, (I - T_1)u \rangle .$$

Remark 3.1.6. The proof below also shows that the quadratic form $\langle T_1 u, (I - T_1)u \rangle \geq 0$.

Proof. The Hermite polynomials with weight $g(v)$ form an orthonormal basis of eigenfunctions for T_1 . Let us denote this by $\{H_n(v)\}_{n=0}^\infty$. Then, since $\langle u, 1 \rangle = 0$, we can write

$$u(v) = \sum_{n \neq 0} c_n H_n(v) .$$

Plugging this expression into the left side of the inequality to be proved,

$$\langle T_1^2 u, (I - T_1)u \rangle = \sum_{n \neq 0} |c_n|^2 \lambda_n^2 (1 - \lambda_n) ,$$

where λ_n are the (non-negative) eigenvalues of T_1 corresponding to H_n . Now, since in the above expression, the sum is over $n \neq 0$, we have that $\lambda_n \leq \frac{1}{2}$ (from Lemma 2.3.3). This yields

$$\begin{aligned} \langle T_1^2 u, (I - T_1)u \rangle &\leq \frac{1}{2} \sum_{n \neq 0} |c_n|^2 \lambda_n (1 - \lambda_n) \\ &= \frac{1}{2} \langle T_1 u, (I - T_1)u \rangle . \end{aligned}$$

□

Lemma 3.1.7. *Let h satisfy the assumptions in Definition 3.1.2. Then*

$$d_1(W_1 h) \leq \frac{2}{\pi} d_1(h) .$$

Proof.

$$\begin{aligned} d_1(W_1 h) &= \sup_{\xi \neq 0} \frac{|f d\theta \hat{h}(\xi \cos \theta) \hat{g}(\xi \sin \theta)|}{|\xi|} \\ &\leq d_1(h) \int d\theta |\cos \theta| \hat{g}(\xi \sin \theta) = d_1(h) \int |\cos \theta| = \frac{2}{\pi} d_1(h) . \end{aligned}$$

□

The next two Lemmas apply, respectively, the previous two Lemmas to bound terms that contain $(1 - \frac{2}{N} + \frac{2}{N}T_1)^m u(v_1)$ and $(1 - \frac{2}{N} + \frac{2}{N}W_1)^m h(v_1)$, which appear in the proof of the Theorem via eq. (61). These remain functions of v_1 alone for any $m \geq 0$.

Lemma 3.1.8. *Under the assumptions of Lemma 3.1.5,*

$$\langle T_1(1 - \frac{2}{N} + \frac{2}{N}T_1)^m u, (I - T_1)(1 - \frac{2}{N} + \frac{2}{N}T_1)^m u \rangle \leq \left(1 - \frac{1}{N}\right)^{2m} \langle T_1 u, (I - T_1)u \rangle .$$

Proof.

$$\begin{aligned} & \langle T_1(1 - \frac{2}{N} + \frac{2}{N}T_1)u, (I - T_1)(1 - \frac{2}{N} + \frac{2}{N}T_1)u \rangle = \\ & (1 - \frac{2}{N})^2 \langle T_1 u, (I - T_1)u \rangle + \frac{4}{N}(1 - \frac{2}{N}) \langle T_1^2 u, (I - T_1)u \rangle + (\frac{2}{N})^2 \langle T_1^3 u, (I - T_1)u \rangle \\ & \leq \left((1 - \frac{2}{N})^2 + \frac{2}{N}(1 - \frac{2}{N}) + \frac{4}{N^2} \frac{1}{4} \right) \langle T_1 u, (I - T_1)u \rangle , \end{aligned}$$

where the last step follows from Lemma 3.1.5. Rearranging the term in the parenthesis above yields $(1 - \frac{1}{N})^2$, and iterating the above result completes the proof. \square

Lemma 3.1.9. *Under the assumptions of Lemma 3.1.7, let $H_m := (1 - \frac{2}{N} + \frac{2}{N}W_1)^m h(v_1)$, $I_m := \mathcal{D}(H_m)$, and $J_m := d_1(H_m)$. Also let $\alpha := 2 - 4/\pi \in (0, 1)$. Then*

- $J_m \leq (1 - \frac{\alpha}{N})^m J_0$.
- $I_m \leq (1 - \frac{1}{N})^m I_0 + \frac{2J_0}{1-\alpha} \left\{ (1 - \frac{\alpha}{N})^m - (1 - \frac{1}{N})^m \right\}$.

Proof. For the first part, we use the convexity of d_1 to write

$$J_m \leq (1 - \frac{2}{N})J_{m-1} + \frac{2}{N}d_1(W_1 H_{m-1}) .$$

Now, by Lemma 3.1.7, we have

$$J_m \leq (1 - \frac{2}{N})J_{m-1} + \frac{2}{N} \frac{2}{\pi} J_{m-1} ,$$

which proves the first part.

For the second part, we start by writing

$$\mathcal{D} \left(\left(1 - \frac{2}{N} + \frac{2}{N} W_1\right) H_{m-1} \right) \leq \left(1 - \frac{2}{N}\right) \mathcal{D}(H_{m-1}) + \frac{2}{N} \mathcal{D}(W_1 H_{m-1}) .$$

Now, we study the term $\mathcal{D}(W_1 H_{m-1})$ as an aside.

$$\begin{aligned} \mathcal{D}(W_1 H_{m-1}) &= \\ \sup_{x_1, x_2 \neq 0} \frac{1}{x_2^2} &\left| \int d\theta \left(\hat{g}(\sqrt{x_1^2 + x_2^2} \cos \theta) H_{m-1}^\wedge(\sqrt{x_1^2 + x_2^2} \sin \theta) - \hat{g}(|x_1| \cos \theta) H_{m-1}^\wedge(|x_1| \sin \theta) \right) \right| \\ &\leq \sup_{x_1, x_2 \neq 0} \int d\theta \frac{\hat{g}(\sqrt{x_1^2 + x_2^2} \cos \theta)}{x_2^2} \left| H_{m-1}^\wedge(\sqrt{x_1^2 + x_2^2} \sin \theta) - H_{m-1}^\wedge(|x_1| \sin \theta) \right| + \\ &\quad \sup_{x_1, x_2 \neq 0} \int d\theta \left| \frac{\hat{g}(\sqrt{x_1^2 + x_2^2} \cos \theta) - \hat{g}(|x_1| \cos \theta)}{x_2^2} \right| |H_{m-1}^\wedge(|x_1| \sin \theta)| \\ &\leq \frac{1}{2} \mathcal{D}(H_{m-1}) + \sup_{x_1, x_2 \neq 0} d_1(H_{m-1}) \int d\theta \hat{g}(|x_1| \cos \theta) \left| \frac{\hat{g}(x_2 \cos \theta) - 1}{x_2^2} \right| |\sin \theta| |x_1| \\ &\leq \frac{1}{2} \mathcal{D}(H_{m-1}) + \nu_g 2\pi^2 d_1(H_{m-1}) \frac{2}{\pi} \sup_{x_1, x_2 \neq 0} \int_0^{\frac{\pi}{2}} d\theta \hat{g}(x_1 \cos \theta) |\sin \theta| |x_1| \cos^2 \theta \\ &\leq \frac{1}{2} \mathcal{D}(H_{m-1}) + d_1(H_{m-1}) . \end{aligned}$$

The last step follows from a computation exactly like that in the proof of Lemma 3.1.4.

Hence, we have

$$\mathcal{D} \left(\left(1 - \frac{2}{N} + \frac{2}{N} W_1\right) H_{m-1} \right) \leq \left(1 - \frac{1}{N}\right) \mathcal{D}(H_{m-1}) + \frac{2}{N} d_1(H_{m-1}) ,$$

which can be written as

$$I_m \leq \left(1 - \frac{1}{N}\right) I_{m-1} + \frac{2}{N} \left(1 - \frac{\alpha}{N}\right)^{m-1} J_0 .$$

To solve the above, set $I_m = c_1 \left(1 - \frac{1}{N}\right)^m I_0 + c_2 \frac{2}{N} \left(1 - \frac{\alpha}{N}\right)^m J_0$. Plugging this into the recurrence relation gives

$$\begin{aligned} c_1 &= 1 - 2 \frac{J_0}{I_0} \frac{1}{1 - \alpha} \\ c_2 &= \frac{N}{1 - \alpha} \end{aligned}$$

From this, the second part of the Lemma is proved. \square

We present two final technical Lemmas, before proceeding to the Theorems.

Lemma 3.1.10. *Let $u(v) \in L^2(\mathbb{R}, g(v)dv)$. Then*

$$\|Qu\| \leq \|u\| ,$$

where the norm is in the space $L^2(\mathbb{R}^N, \gamma(\mathbf{v})d\mathbf{v})$.

Proof. This follows directly from Lemma 2.3.1. □

Lemma 3.1.11. *Let h satisfy the assumptions in Definition 3.1.2. Then*

$$\sup_{\vec{\xi} \neq 0} \frac{|\widehat{Qh}(\vec{\xi})|}{|\vec{\xi}|^2} \leq \sup_{\vec{\xi} \neq 0} \frac{|\hat{h}(\vec{\xi})|}{|\vec{\xi}|^2} .$$

Proof. This follows directly from Lemma 2.5.6. □

Theorem 3.1.12. *Let $f(v)$ in $L^1(\mathbb{R})$ be a probability density such that $u(v) := \frac{f(v)}{g(v)} \in L^2(\mathbb{R}, g(v)dv)$. Then*

$$\|e^{Nt(Q-I)}u - e^{2t(T-I)}u\| \leq \frac{\sqrt{2}}{\sqrt{N-1}} \|u - 1\| (1 - e^{-t}) ,$$

where the above norm is in $L^2(\mathbb{R}^N, \gamma(\mathbf{v})d\mathbf{v})$.

Proof. Starting with eq. (61) and using the sub-linearity of the norm, and Lemmas 3.1.10, 3.1.1 and 3.1.8,

$$\|e^{Nt(Q-I)}u - e^{2t(T-I)}u\| \leq \frac{e^{-Nt}}{\sqrt{N-1}} \sum_{k=0}^{\infty} \frac{N^k t^k}{k!} \left(\frac{2}{N}\right) \sum_{m=0}^{k-1} \left(1 - \frac{1}{N}\right)^m \sqrt{\langle T_1(u-1), (I-T_1)(u-1) \rangle} .$$

The assumptions of Lemma 3.1.8 are satisfied by $u - 1$, which we introduce through the observation that $\langle T_1 u, (I - T_1)u \rangle = \langle T_1(u-1), (I - T_1)(u-1) \rangle$.

$$\begin{aligned} &= \frac{e^{-Nt}}{\sqrt{N-1}} \left(\frac{2}{N}\right) \sqrt{\langle T_1(u-1), (I-T_1)(u-1) \rangle} \sum_{k=0}^{\infty} \frac{N^k t^k}{k!} \frac{1 - (1 - \frac{1}{N})^k}{\frac{1}{N}} \\ &= \frac{2e^{-Nt}}{\sqrt{N-1}} \sqrt{\langle T_1(u-1), (I-T_1)(u-1) \rangle} \sum_{k=0}^{\infty} \frac{t^k}{k!} (N^k - (N-1)^k) \\ &= \frac{2\sqrt{\langle T_1(u-1), (I-T_1)(u-1) \rangle}}{\sqrt{N-1}} (1 - e^{-t}) . \end{aligned}$$

Noting that $\sqrt{\langle T_1(u-1), (I-T_1)(u-1) \rangle} \leq \frac{1}{\sqrt{2}} \|u - 1\|$, we have the proof. □

Theorem 3.1.13. Let $f(v) \in L^1$ such that $\int f(v)dv = 1$ and $\int f(v)v^2dv =: \nu_f < \infty$.

Then we have

$$d_2 \left(e^{N(Q-I)t} f(v_1) \prod_{j=2}^N g(v_j), e^{2(W_1-I)t} f(v_1) \prod_{j=2}^N g(v_j) \right) \leq \frac{1}{N-1} \left\{ \left[I_0 - \frac{2J_0}{1-\alpha} \right] (1 - e^{-t}) + \left[\frac{J_0}{1-\alpha} + J_0 \right] (1 - e^{-\alpha t}) \right\},$$

where $I_0 = \mathcal{D}((f-g)_e)$, $J_0 = d_1((f-g)_e)$.

Remark. The second moment assumption on f guarantees that I_0 and J_0 are finite.

Proof. Let $h(v) = f(v) - g(v)$. Note that we can replace f by h in the above expression. Moreover, since the action of operators $N(Q-I)$ and $2(W_1-I)$ on the odd part of h is exactly the same, h can be replaced by its even part. This observation, coupled with the assumption that $\nu_f < \infty$ ensure that the even part of h satisfies the assumptions in Definition 3.1.2. For the remainder of the proof, we assume without loss of generality that h is the even part of $f - g$.

Starting as in the L^2 case from eq. (61) and using the linearity of the Fourier transform,

$$d_2 \left(e^{N(Q-I)t} f(v_1) \prod_{j=2}^N g(v_j), e^{2(W_1-I)t} f(v_1) \prod_{j=2}^N g(v_j) \right) \leq e^{-Nt} \sum_{k=0}^{\infty} \frac{t^k}{k!} N^k \left(\frac{2}{N} \right) \sum_{m=0}^{k-1} \sup_{\vec{\xi} \neq \mathbf{0}} \frac{\left| \mathcal{F} \left(Q^{k-m-1} \left(\frac{1}{N-1} \sum_{i=2}^N R_{1i} - W_1 \right) \left(1 - \frac{2}{N} + \frac{2}{N} W_1 \right)^m h(v_1) \prod_{j=2}^N g(v_j) \right) \right|}{|\vec{\xi}|^2}.$$

Now, recalling that $H_m := \left(1 - \frac{2}{N} + \frac{2}{N} W_1 \right)^m h(v_1)$, and using Lemma 3.1.11,

$$d_2 \left(e^{N(Q-I)t} f(v_1) \prod_{j=2}^N g(v_j), e^{2(W_1-I)t} f(v_1) \prod_{j=2}^N g(v_j) \right) \leq e^{-Nt} \sum_{k=0}^{\infty} \frac{t^k}{k!} N^k \left(\frac{2}{N} \right) \sum_{m=0}^{k-1} d_2 \left(\frac{1}{N-1} \sum_{i=2}^N R_{1i} H_m(v_1) \prod_{j=2}^N g(v_j), W_1 H_m(v_1) \prod_{j=2}^N g(v_j) \right) e^{-Nt} \sum_{k=0}^{\infty} \frac{t^k}{k!} N^k \left(\frac{2}{N} \right) \sum_{m=0}^{k-1} \frac{1}{N-1} \left[\frac{1}{2} \mathcal{D}(H_m) + d_1(H_m) \right],$$

where the last step follows from Lemma 3.1.4. Now, finally, we apply Lemma 3.1.9 to complete the proof.

$$d_2 \left(e^{N(Q-I)t} f(v_1) \prod_{j=2}^N g(v_j), e^{2(W_1-I)t} f(v_1) \prod_{j=2}^N g(v_j) \right) \leq$$

$$\frac{2e^{-Nt}}{N-1} \sum_{k=0}^{\infty} \frac{N^k t^k}{k!} \left(\left(\frac{I_0}{2} - \frac{J_0}{1-\alpha} \right) \left(1 - \left(1 - \frac{1}{N} \right)^k \right) + \left(\frac{J_0}{1-\alpha} + J_0 \right) \left(1 - \left(1 - \frac{\alpha}{N} \right)^k \right) \right).$$

Summing the series, we get the desired result. \square

In the next section, we look at another route to realizing the thermostat model W : as a suitably defined van Hove limit of a partially thermostated system.

3.2 *As a van Hove Limit*

Here, we start with Model II when there are only two particles ($N = 2$, $m = 1$) described by eq. (12):

$$\frac{\partial f^\lambda}{\partial t} = -2\lambda(I - R_{12})f^\lambda - \mu(I - S_1)f^\lambda =: -\mathcal{G}^\lambda f^\lambda. \quad (66)$$

Here the superscript makes it explicit that the solution depends on λ .

Particle 2 interacts through the Kac collision with Particle 1, which is given the Gaussian distribution $g(v) = \sqrt{\frac{1}{2\pi}} e^{-\frac{v^2}{2}}$ at random times due to the action of the strong thermostat S_1 . We increase the rate μ at which this acts relative to the rate of the Kac collision 2λ . This can be achieved by increasing the time scale of the Kac operator $\frac{1}{2\lambda} \rightarrow \infty$ and sampling at longer time intervals $\tau := t\lambda$. Thus, the strong thermostat, operating on a much smaller time-scale, becomes powerful in the limit. The result is that by passing through a van Hove (weak-coupling, large time) limit [12] of this system, Particle 2 gets thermostated, via its interaction with Particle 1 whose distribution is essentially always $g(v)$.

We are interested in the evolution of $\tilde{f}^\lambda(v_1, v_2, \tau) := f^\lambda(v_1, v_2, \frac{\tau}{\lambda})$ in the limit $\lambda \rightarrow 0$. Here $f^\lambda(v_1, v_2, t)$ satisfies (66) above. The equation satisfied by $\tilde{f}^\lambda(v_1, v_2, \tau)$

is then:

$$\frac{\partial \tilde{f}^\lambda}{\partial \tau} = -2(I - R_{12})\tilde{f}^\lambda - \frac{\mu}{\lambda}(I - S_1)\tilde{f}^\lambda =: -\frac{\mathcal{G}^\lambda}{\lambda}\tilde{f}^\lambda. \quad (67)$$

We have the following theorem, which states that the diagram in Figure 1 commutes.

Theorem 3.2.1. *Let \tilde{f}^λ satisfy eq. (67) with initial condition $\tilde{f}^\lambda(v_1, v_2, 0) = \phi(v_1, v_2) \in L^1(\mathbb{R}^2)$. Then for $\tau > 0$, $\lim_{\lambda \rightarrow 0} \tilde{f}^\lambda =: g(v_1)\tilde{f}(v_2, \tau)$ exists in $L^1(\mathbb{R}^2)$, where \tilde{f} satisfies the equation*

$$\frac{\partial \tilde{f}}{\partial \tau} = -2(I - W_2)\tilde{f} \quad (68)$$

together with the initial condition $\tilde{f}(v_2, 0) = \frac{S_1 \phi(v_1, v_2)}{g(v_1)}$. W_2 is the thermostat (9) acting on v_2 .

$$\begin{array}{ccc} \phi(v_1, v_2) & \xrightarrow{e^{-\frac{\tau}{\lambda}\mathcal{G}^\lambda}} & \tilde{f}^\lambda(v_1, v_2, \tau) \\ S_1 \downarrow & & \downarrow \lambda \rightarrow 0 \\ g(v_1)\tilde{f}(v_2, 0) & \xrightarrow{e^{-2\tau(I-W_2)}} & g(v_1)\tilde{f}(v_2, \tau) \end{array}$$

Figure 1: van Hove Limit

Proof. We can write $e^{-\frac{\mu}{\lambda}\tau(I-S_1)} = I + (I - S_1)(e^{-\mu\tau/\lambda} - 1)$ because $(I - S_1)$ is idempotent. This implies that

$$\|e^{-\frac{\mu}{\lambda}\tau(I-S_1)} - S_1\|_1 = e^{-\mu\tau/\lambda}\|I - S_1\|_1 \leq 2e^{-\mu\tau/\lambda}. \quad (69)$$

For each λ , the operators in $\frac{1}{\lambda}\mathcal{G}^\lambda$ are bounded. Thus, the Dyson expansion (the infinite series version of the Duhamel formula) corresponding to the evolution in (67) gives $e^{-\frac{\tau}{\lambda}\mathcal{G}^\lambda}\phi = \sum_{k=0}^{\infty} b_k(\phi)$ where

$$b_0(\phi) = e^{-\frac{\mu}{\lambda}(I-S_1)\tau}\phi,$$

$$b_1(\phi) = \int_{t_1=0}^{\tau} e^{-\frac{\mu}{\lambda}(I-S_1)(\tau-t_1)} [-2(I-R_{12})] e^{-\frac{\mu}{\lambda}(I-S_1)t_1} \phi \, dt_1, \text{ and}$$

$$b_k(\phi) = \int_{\{0 \leq t_k \leq \dots t_1 \leq \tau\}} e^{-\frac{\mu}{\lambda}(I-S_1)(\tau-t_1)} [-2(I-R_{12})] e^{-\frac{\mu}{\lambda}(I-S_1)(t_1-t_2)} \dots [-2(I-R_{12})] e^{-\frac{\mu}{\lambda}(I-S_1)t_k} \phi \, d\vec{t}$$

Using (69) and the identity $S_1 R_{12} S_1 = S_1 W_2 = W_2 S_1$, we show that $\forall k$, $b_k(\phi)$ converges to

$$\int_{\{0 \leq t_k \leq \dots t_1 \leq \tau\}} S_1 [-2(I-R_{12})] S_1 \dots [-2(I-R_{12})] S_1 \phi \, d\vec{t} = \frac{1}{k!} (-2(I-W_2))^k (S_1 \phi)$$

in L^1 as $\lambda \rightarrow 0$.

Finally, we use the fact that for each $u \geq 0$, $\|e^{-\frac{\mu}{\lambda}(I-S_1)u} \phi\|_1 = \|\phi\|_1$ and $\|(I-R_{12})\phi\|_1 \leq 2\|\phi\|_1$ so that $\|b_k(\phi)\| \leq 4^k \int_{\{0 \leq t_k \leq \dots t_1 \leq \tau\}} dt_1 \dots dt_k \|\phi\|_1 = \frac{(4\tau)^k}{k!} \|\phi\|_1$, independently of λ . Therefore the dominated convergence theorem can be applied to give

$$\begin{aligned} \lim_{\lambda \rightarrow 0} e^{-\frac{\tau}{\lambda} \mathcal{G}^\lambda} \phi &= \lim_{\lambda \rightarrow 0} \sum_{k=0}^{\infty} b_k(\phi) = \sum_{k=0}^{\infty} \lim_{\lambda \rightarrow 0} b_k(\phi) \\ &= \sum_{k=0}^{\infty} (-2(I-W_2))^k \frac{\tau^k}{k!} (S_1 \phi) = e^{-2(I-W_2)\tau} (S_1 \phi). \end{aligned}$$

□

Given this connection between the strong thermostat S and the thermostat W , it is worth revisiting the entropy bound (22) for Model II. Upon making the transformation $(\mu, \lambda) \rightarrow (\frac{\mu}{\lambda}, 1)$ corresponding to the van Hove limit (see eq. (67)), we obtain the bound

$$S(t) \leq e^{-t} S(0)$$

as $\lambda \rightarrow 0$. This is exactly the optimal entropy production bound (43) for the thermostat W (Note: the thermostat here appears with a factor of 2, owing to the 2λ term).

CHAPTER IV

CONCLUSION AND OUTLOOK

There is still much distance to cover en route to understanding the non-equilibrium behavior of the Kac model, and extrapolating this to gain insight into general non-equilibrium systems. At this point, we have an understanding of physically relevant quantities like the spectral gap and entropy decay rates. For Model I, the first and second gaps were computed and the entropic convergence to equilibrium, and equilibration in the GTW metric were established in a quantitative fashion. Here, the approach to equilibrium persisted uniformly in N . Moreover, since propagation of chaos holds, a rigorous connection was made with a Boltzmann-type equation. It is conceivable that a propagation of chaos result *uniform in time* can be obtained using methods similar to those used in [27] for the isolated Kac model.

Moreover, Model II, the partially thermostated system, provides clues to the scaling of equilibration time-scales, and raises the question of whether “eventually exponential” entropy decay is a possible occurrence in systems like these. Our results imply that if a macroscopic fraction of particles is thermostated, the kinetic energy and the L^2 norm decay exponentially to their respective equilibrium values at a rate independent of N . However, our entropy bound and the GTW distance yields a decay rate that vanishes as $N \rightarrow \infty$ in the thermodynamic limit. Hence, at least under a suitable class of initial conditions, we think it should be possible to improve (22) to reflect the physical situation. The question of entropy production at $t = 0$ (and any N) remains unsettled. The bound (22) does not preclude the possibility of zero entropy production at time 0. However, we do not know if it actually occurs in the model for some initial conditions.

We could also investigate the case of Model II where the strong thermostat S is replaced by the more physical thermostat W . We believe that the qualitative nature of the results will be similar, but the proofs will be slightly more complicated. On the other hand, Model I with the thermostat W replaced by S is simple to analyze, given our results. For example, the spectral gap and the exponential decay rate in entropy in this case become μ . It is interesting to note that this system is the $m = N$ case of Model II, and so one could wonder if it is possible to improve the entropy bound (Theorem 2.2.5) for Model II to obtain one that tends to $e^{-\mu t}$ as $m \nearrow N$.

A question that follows naturally from Chapter 3 is whether the fully thermostated Model I can be realized as a large-size limit of systems with non-ideal heat baths. In this case, the scaling must ensure that the Kac collisions among the particles survive in the limit.

Apart from the questions posed above to further the understanding of the results and to close some gaps, there are some allied problems that are interesting. First, we would like to be able to generalize our analysis on these problems to the momentum-preserving three-dimensional Kac collision [8]. Next, we could consider the sustained non-equilibrium situation and ask what the steady-state looks like when we couple a Kac system to two unequal heat baths. A quick check tells us that the steady-state is *not* a Gaussian. Numerical explorations could be a possible approach to finding the steady-state. Also, it would be of interest to study large systems in which spatially inhomogeneity is intertwined with the equilibration process. One step to this end is to consider a network of Kac systems with unequal heat baths at some nodes, and interactions with neighboring nodes alone.

APPENDIX A

SPECTRUM OF EVOLUTION OPERATOR FOR

$$N = 2, M = 1$$

We analyze the spectrum of the self-adjoint evolution operator $\mathcal{L}_{2,1} = 2\lambda(I - Q) + \mu(I - P_1)$, in the space $L^2(\mathbb{R}^2, \gamma(\mathbf{v})d\mathbf{v})$, and deduce its spectral gap stated in (36). For simplicity, we denote the operators $\mathcal{L}_{2,1}$ and P_1 by \mathcal{L} and P .

Notice that \mathcal{L} is a linear combination of two projections ($Q \equiv R_{12}$ is an orthogonal projection onto radial functions in \mathbb{R}^2). The condition $\langle h, 1 \rangle = 0$, corresponding to the normalization of $f = \gamma(1 + h)$, leads us to work in the space of Hermite polynomials $\{H_\alpha(v)\}_{\alpha=0}^\infty$ with weight $g(v)$. The space of interest \mathcal{X}_2 is spanned by $\{K_{i,j} : i, j \in \mathbb{N}, (i, j) \neq (0, 0)\}$, where $K_{i,j} := H_i(v_1)H_j(v_2)$. Without loss of generality, we work with monic Hermite polynomials.

The action of P is as follows:

$$PK_{i,j} = \begin{cases} 0 & : i \neq 0 \\ K_{0,j} & : i = 0 \end{cases}$$

Since each term in $K_{i,j}$ is odd in either v_1 or v_2 when either i or j is odd, we have that $QK_{i,j} = 0$ when either i or j is odd. We deduce the action of Q on $K_{2\alpha_1, 2\alpha_2}$ from its action on $v_1^{2\alpha_1}v_2^{2\alpha_2}$ using the following Lemma from Ref. [3], which applies to Q as it is a projection onto radial functions.

Lemma A.0.2 (Ref. [3]). *Let A be a self-adjoint operator on $L^2(\mathbb{R}^N, \gamma(\mathbf{v})d\mathbf{v})$ that preserves the space P_{2l} of homogeneous even polynomials in v_1, \dots, v_N of degree $2l$. If*

$$A(v_1^{2\alpha_1} \dots v_N^{2\alpha_N}) = \sum_{\sum \alpha_i = \sum \beta_i} c_{\beta_1 \dots \beta_N} v_1^{2\beta_1} \dots v_N^{2\beta_N},$$

we get

$$A(H_{2\alpha_1}(v_1)\dots H_{2\alpha_N}(v_N)) = \sum_{\sum \alpha_i = \sum \beta_i} c_{\beta_1\dots\beta_N} H_{2\beta_1}(v_1)\dots H_{2\beta_N}(v_N) .$$

Let $n := \alpha_1 + \alpha_2$ and $\Gamma_{\alpha_1, \alpha_2} := \int_0^{2\pi} \cos^{2\alpha_1} \theta \sin^{2\alpha_2} \theta d\theta = \frac{(2\alpha_1-1)!!(2\alpha_2-1)!!}{2^{\alpha_1+\alpha_2}(\alpha_1+\alpha_2)!}$, with the standard definition $(-1)!! = 1$. Then we have

$$QK_{i,j} = \begin{cases} 0 & : i \text{ or } j \text{ odd} \\ \Gamma_{\alpha_1, \alpha_2} \sum_{m=0}^n \binom{n}{m} K_{2m, 2n-2m} & : i = 2\alpha_1, j = 2\alpha_2 \end{cases}$$

Now a case-by-case analysis, using the fact that $L_{2n} := \text{Span}\{H_{2\alpha_1}(v_1)H_{2\alpha_2}(v_2) : \alpha_1 + \alpha_2 = n\}$ are invariant subspaces for \mathcal{L} , yields the following for the spectrum of \mathcal{L} :

Eigenvalue	Eigenfunction
$2\lambda + \mu$	$K_{i,j}$, i or j odd, $i \neq 0$ $\sum_{i=1}^n c_i K_{2i, 2n-2i}$ where $\sum_{i=1}^n c_i \Gamma_{i, n-i} = 0$
2λ	$K_{0,j}$, j odd
$x^{\pm, n}$	$\sum_{i=0}^n c_i^{\pm, n} K_{2i, 2n-2i}$ and eq. (70)

Remark A.0.3. The first row corresponds to functions that belong to the kernels of both Q and P , and the second row to functions that belong to the kernels of Q and $I - P$.

Here,

$$x^{\pm, n} = \frac{(2\lambda + \mu) \pm \sqrt{(2\lambda + \mu)^2 - 8\lambda\mu(1 - \Gamma_{0,n})}}{2}$$

and

$$c_0^{\pm, n} = \frac{2\lambda}{2\lambda - x^{\pm, n}} \text{ and } c_i^{\pm, n} = \frac{2\lambda \binom{n}{i}}{x^{\mp, n}} \text{ for } i \neq 0 . \quad (70)$$

Using the fact that $\Gamma_{0,n} = \frac{1}{2\pi} \int_0^{2\pi} \cos^{2n} \theta d\theta$ is decreasing in n , it is easy to see that the smallest eigenvalue is $x^{-,1}$. The corresponding eigenfunction is $\frac{2\lambda}{2\lambda - x^{-,1}} K_{0,2} + \frac{2\lambda}{x^{+,1}} K_{2,0}$.

APPENDIX B

OPTIMIZING SEQUENCE FOR INEQUALITY (56)

Our aim is to find a sequence of densities h_ϵ on $L^1(\mathbb{R}^2, \gamma d\mathbf{v})$ such that

$$\frac{S(Qh_\epsilon) + S(P_1h_\epsilon)}{S(h_\epsilon)} \geq 2 - \epsilon$$

for any $\epsilon > 0$.

Choose $h = \frac{1}{Z_a} \mathbf{1}_{[-a,a]} \frac{1}{Z_a^R} \mathbf{1}_{[R-a, R+a]}$. Here $Z_a = \int_{-a}^a \sqrt{\frac{\beta}{2\pi}} e^{-\frac{\beta v^2}{2}} dv$ and $Z_a^R = \int_{R-a}^{R+a} \sqrt{\frac{\beta}{2\pi}} e^{-\frac{\beta v^2}{2}} dv$.

We have $S(h) = -\log(Z_a Z_a^R)$.

Then $P_1 h = \frac{1}{Z_a^R} \mathbf{1}_{[R-a, R+a]}$ so that $S(P_1 h) = -\log(Z_a^R)$.

For the entropy of Qh , we bypass the exact computation and instead use the fact that Qh is a density supported on the annular region Γ with inner and outer radii $R_1 = R - a$ and $R_2 = \sqrt{(R + a)^2 + a^2}$. Then the entropy of the function $K(v_1, v_2) := \frac{1}{Z_\Gamma} \mathbf{1}_\Gamma$ (Z_Γ is the normalization) bounds from below the entropy of the function Qh . That is,

$$S(K) \leq S(Qh).$$

(The above follows for instance, from the observation that $\int_\Gamma \frac{Qh}{(1/Z_\Gamma)} \log \left(\frac{Qh}{(1/Z_\Gamma)} \right) \gamma d\mathbf{v} \geq 0$.)

Note that $S(K) = -\log Z_\Gamma$. Therefore, we have

$$\frac{S(P_1 h) + S(Qh)}{S(h)} \geq \frac{S(P_1 h) + S(K)}{S(h)} = \frac{-\log Z_a^R - \log Z_\Gamma}{-\log Z_a - \log Z_a^R}.$$

Recall that Z_a is independent of R . Let us study the dependency of the other terms on R . First,

$$Z_a^R = \int_{R-a}^{R+a} \sqrt{\frac{\beta}{2\pi}} e^{-\frac{\beta v^2}{2}} dv = \int_{-a}^a \sqrt{\frac{\beta}{2\pi}} e^{-\frac{\beta w^2}{2}} e^{-\frac{\beta R^2}{2}} e^{-\beta w R} dw$$

or

$$Z_a^R \leq e^{-\frac{\beta R^2}{2}} e^{\beta R a} \int_{-a}^a \sqrt{\frac{\beta}{2\pi}} e^{-\frac{\beta w^2}{2}} dw .$$

Also,

$$Z_\Gamma = \int_{R_1}^{R_2} \beta e^{-\beta/2 r^2} r dr = e^{-\frac{\beta}{2}(R-a)^2} - e^{-\frac{\beta}{2}[(R+a)^2+a^2]} .$$

Finally, we get that

$$\frac{S(P_1 h) + S(Q h)}{S(h)} \geq \frac{-\log \left(e^{-\frac{\beta}{2} R^2} e^{\beta R a} Z_a \right) - \log \left(e^{-\frac{\beta}{2}(R-a)^2} - e^{-\frac{\beta}{2}[(R+a)^2+a^2]} \right)}{-\log Z_a - \log \left(e^{-\frac{\beta}{2} R^2} e^{\beta R a} Z_a \right)} ,$$

where the right-hand-side goes to 2 as $R \rightarrow \infty$.

REFERENCES

- [1] BAKRY, D. and ÉMERY, M., “Diffusions hypercontractives,” in *Séminaire de probabilités, XIX, 1983/84*, vol. 1123 of *Lecture Notes in Math.*, pp. 177–206, Berlin: Springer, 1985.
- [2] BOBYLEV, A. and CERCIGNANI, C., “On the rate of entropy production for the Boltzmann equation,” *J. Stat. Phys.*, vol. 94, no. 3-4, pp. 603–618, 1999.
- [3] BONETTO, F., LOSS, M., and VAIDYANATHAN, R., “The Kac model coupled to a thermostat,” *J. Stat. Phys.*, vol. 156, no. 4, pp. 647–667, 2014.
- [4] CALLEN, H. B., *Thermodynamics*. New York, N.Y.: John Wiley & Sons, Inc., 1960.
- [5] CARLEN, E. A., “The rate of local equilibration in kinetic theory,” in *Prospects in mathematical physics*, vol. 437 of *Contemp. Math.*, pp. 71–88, Providence, RI: Amer. Math. Soc., 2007.
- [6] CARLEN, E. A., CARVALHO, M. C., LOSS, M., ROUX, J. L., and VILLANI, C., “Entropy and chaos in the Kac model,” *Kinet. Relat. Models*, vol. 3, pp. 85–122, 2010.
- [7] CARLEN, E. A., CARVALHO, M. C., and LOSS, M., “Many-body aspects of approach to equilibrium,” in *Journées “Équations aux Dérivées Partielles” (La Chapelle sur Erdre, 2000)*, pp. Exp. No. XI, 12, Nantes: Univ. Nantes, 2000.
- [8] CARLEN, E. A., GERONIMO, J. S., and LOSS, M., “Determination of the spectral gap in the Kac model for physical momentum and energy-conserving collisions,” *SIAM J. Math. Anal.*, vol. 40, no. 1, pp. 327–364, 2008.
- [9] CARLEN, E. A., LEBOWITZ, J. L., and MOUHOT, C., “Exponential approach to, and properties of, a non-equilibrium steady state in a dilute gas,” *arXiv:1406.4097*.
- [10] CERCIGNANI, C., “H-theorem and trend to equilibrium in the kinetic theory of gases,” *Archiv of Mechanics, Archiwum Mechaniki Stosowanej*, vol. 34, pp. 231–241, 1982.
- [11] CHURCHILL, R. V., *Operational Mathematics, 3rd ed.* McGraw Hill, New York, 1972.
- [12] DELL’ANTONIO, G., *The van Hove limit in classical and quantum mechanics*, pp. 75–110. Springer Berlin / Heidelberg, 1982.

- [13] EINAV, A., “On Villani’s conjecture concerning entropy production for the Kac master equation,” *Kinet. Relat. Models*, vol. 4, no. 2, pp. 479–497, 2011.
- [14] GABETTA, G., TOSCANI, G., and WENNBORG, B., “Metrics for probability distributions and the trend to equilibrium for solutions of the boltzmann equation,” *Journal of Statistical Physics*, vol. 81, no. 5-6, pp. 901–934, 1995.
- [15] GALLAVOTTI, G., *Statistical Mechanics: A Short Treatise (Theoretical and Mathematical Physics)*. Springer, softcover reprint of hardcover 1st ed. 1999 ed., 11 2010.
- [16] GROSS, L., “Logarithmic Sobolev inequalities,” *Amer. J. Math.*, vol. 97, no. 4, pp. 1061–1083, 1975.
- [17] GROSS, L., “Logarithmic Sobolev inequalities and contractivity properties of semigroups,” in *Dirichlet forms (Varenna, 1992)*, vol. 1563 of *Lecture Notes in Math.*, pp. 54–88, Berlin: Springer, 1993.
- [18] HAN, T. S., “Nonnegative entropy measures of multivariate symmetric correlations,” *Information and Control*, vol. 36, no. 2, pp. 133–156, 1978.
- [19] ILLNER, R. and PULVIRENTI, M., “Global validity of the Boltzmann equation for two- and three-dimensional rare gas in vacuum. Erratum and improved result: “Global validity of the Boltzmann equation for a two-dimensional rare gas in vacuum” [Comm. Math. Phys. **105** (1986), no. 2, 189–203; MR0849204 (88d:82061)] and “Global validity of the Boltzmann equation for a three-dimensional rare gas in vacuum” [ibid. **113** (1987), no. 1, 79–85; MR0918406 (89b:82052)] by Pulvirenti,” *Comm. Math. Phys.*, vol. 121, no. 1, pp. 143–146, 1989.
- [20] JANVRESSE, E., “Spectral gap for Kac’s model of Boltzmann equation,” *Ann. Probab.*, vol. 29, no. 1, pp. 288–304, 2001.
- [21] KAC, M., “Foundations of kinetic theory,” in *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III*, (Berkeley and Los Angeles), pp. 171–197, University of California Press, 1956.
- [22] LANFORD, III, O. E., “Time evolution of large classical systems,” in *Dynamical systems, theory and applications (Recontres, Battelle Res. Inst., Seattle, Wash., 1974)*, pp. 1–111. Lecture Notes in Phys., Vol. 38, Berlin: Springer, 1975.
- [23] LANFORD, III, O. E., “On a derivation of the Boltzmann equation,” in *International Conference on Dynamical Systems in Mathematical Physics (Rennes, 1975)*, pp. 117–137. Astérisque, No. 40, Paris: Soc. Math. France, 1976.
- [24] LOOMIS, L. H. and WHITNEY, H., “An inequality related to the isoperimetric inequality,” *Bull. Amer. Math. Soc.*, vol. 55, pp. 961–962, 1949.
- [25] MASLEN, D. K., “The eigenvalues of Kac’s master equation,” *Math. Z.*, vol. 243, no. 2, pp. 291–331, 2003.

- [26] MCKEAN, JR., H. P., “Speed of approach to equilibrium for Kac’s caricature of a Maxwellian gas,” *Arch. Rational Mech. Anal.*, vol. 21, pp. 343–367, 1966.
- [27] MISCHLER, S. and MOUHOT, C., “Kac’s program in kinetic theory,” *Invent. Math.*, vol. 193, no. 1, pp. 1–147, 2013.
- [28] STAM, A. J., “Some inequalities satisfied by the quantities of information of Fisher and Shannon,” *Information and Control*, vol. 2, pp. 101–112, 1959.
- [29] TOSCANI, G., “Entropy production and the rate of convergence to equilibrium for the Fokker-Planck equation,” *Quart. Appl. Math.*, vol. 57, no. 3, pp. 521–541, 1999.
- [30] TOSSOUNIAN, H. and VAIDYANATHAN, R., “Partially thermostated kac model,” *Journal of Mathematical Physics*, vol. 56, no. 8, pp. –, 2015.
- [31] VILLANI, C., “Cercignani’s conjecture is sometimes true and always almost true,” *Comm. Math. Phys.*, vol. 234, no. 3, pp. 455–490, 2003.